Dual Set Membership Filter With Minimizing Nonlinear Transformation of Ellipsoid

Zhiguo Wang, Xiaojing Shen, Haiqi Liu, Fanqin Meng, and Yunmin Zhu

***Abstract*—In this article, we propose a dual set membership filter for nonlinear dynamic systems with additive unknown but bounded noises, and it has three distinct advantages. First, the nonlinear system is translated into the linear system by leveraging a semi-infinite programming, rather than linearizing the nonlinear function. The semi-infinite programming is to find an ellipsoid bounding the nonlinear transformation of an ellipsoid, which aims to compute a tight ellipsoid to cover the state. Second, the duality result of the semi-infinite programming is derived by rigorous analysis; then, a first-order Frank–Wolfe method is developed to efficiently solve it with a lower computation complexity. Third, the proposed filter enjoys stability for some special nonlinear dynamic systems and succeeds the advantages of the classic linear set membership filter. Finally, two illustrative examples in the simulations reveal the effectiveness of the dual set membership filter.**

***Index Terms*—Frank–Wolfe (FW) method, mobile robot localization, nonlinear system, set membership filter (SMF), unknown but bounded noise.**

I. Introduction

THE nonlinear filter is an important research problem in many fields, such as target tracking [1], navigation [2], and mobile robot [3]. It is extensively researched in the Bayesian framework, which is based on stochastic assumptions about the process and measurement noises. If the exact distribution of these noises is known, then the classic nonlinear Bayesian filter, such as extended Kalman filter [4], unscented Kalman filter (UKF) [5], and particle filter [6], can obtain the better estimation. However, the statistical properties of the process noise and measurement noise may be imprecisely known, which can degrade performance of the Bayesian filter. It then seems more natural to assume that the state perturbations and measurement noise are unknown but bounded [7].

In this article, we consider the problem of ellipsoidal set membership filter (SMF) for nonlinear dynamic systems with additive unknown but bounded noises, which does not require any assumption on the noise statistics. The SMF for linear systems was first proposed by Schweppe [8], and its basic idea is propagating bounding ellipsoids [9] (or polyhedron, boxes, and zonotopes) for dynamic systems with bounded noises. Recently, the SMF has also been extensively explored (see [10]–[14] and references therein).

It is difficult to extend the SMF to nonlinear dynamic systems, especially for online implementation. The reason is that a general nonlinear function maps an ellipsoid to an irregular set, which makes the set operation very complicated. Some researchers have considered the SMF for nonlinear systems [15]–[19]. In [20], Scholte and Campbell develop the extended set membership filter (ESMF) for a general class of nonlinear systems with online usage. Specifically, the nonlinear dynamics are linearized about the current estimate; then, the higher order remainder terms are then bounded by interval mathematics [21], and the remainder bounds are incorporated as additions to the process or sensor noise bounds; finally, the classis SMF for a linear system can be used. In [22], Yang and Li employ the fuzzy modeling approach to approximate the nonlinear systems and the S-procedure technique to determine a state estimation ellipsoid. In [23], Calafiore proposes a nonlinear set membership filter (NSMF) based on a two-step prediction–correction. Each step of this filter requires to solve a semidefinite optimization problem (SDP) to obtain the optimal outer-bounding ellipsoid.

The limitation of current SMF approaches for the nonlinear system is that it needs to linearize the nonlinear function. However, linearized transformations are only reliable if the nonlinear function can be well approximated by a linear function [5]. Otherwise, for a general nonlinear function, the linearized approximation can be extremely poor [24], which leads to a bigger ellipsoid to bound the higher order remainder. In fact, the linearization is to find an ellipsoid to approximate the irregular set, which is the nonlinear transformation of an ellipsoid. If we can directly calculate a minimum size ellipsoid to cover the irregular set, the performance of the SMF can be better (see Fig. 1). Thus, these facts motivate us to consider the NSMF without linearizing the nonlinear functions.

On the other hand, the current SMF is that each step needs to solve an SDP problem [12], [25]; of course, it can be solved by an interior point method. However, as problems have grown in size, such as large-scale problems in multiple target tracking, it does not work well due to the higher computation complexity [26]. Hence, people have renewed their interest in the first-order method, which can be well extended to such a

Manuscript received February 19, 2020; revised November 10, 2020; accepted May 8, 2021. Date of publication May 18, 2021; date of current version April 26, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61673282 and Grant PCSIRT16R53. Recommended by Associate Editor L. Zhang. *(Corresponding author: Xiaojing Shen.)*

The authors are with the College of Mathematics, Sichuan University, Chengdu 610064, China (e-mail: wangzg315@126.com; xiao23332@163.com; 411566818@qq.com; 1076019964@qq.com; ymzhu@scu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TAC.2021.3081078.

Digital Object Identifier 10.1109/TAC.2021.3081078

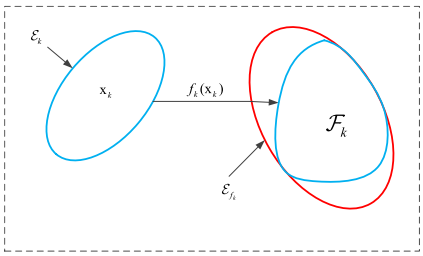

Fig. 1. Illustration of the optimization problem (12), (13).

large-scale problem. In addition, the Frank–Wolfe (FW) method is a popular first-order method, and it can efficiently solve the minimum-volume enclosing ellipsoid problem for the convex hull of a finite set of points [27]. In this article, we consider developing the FW method to the SMF, which makes each step of the SMF very cheap to perform.

The main contributions of this article are as follows. Based on the advanced optimization technique, we propose a dual set membership filter (DSMF) to recursively compute a bounding ellipsoid to cover the state, and it has three vital edges.

1) Compared with the traditional nonlinear SMF, we use the semi-infinite programming to translate the nonlinear system into the linear system, rather than linearizing the nonlinear function, which can obtain a tighter ellipsoid that contains the true state.
2) By rigorous analysis, the dual problem of the semi-infinite programming is derived. Moreover, for online usage, we propose the first-order FW method to solve it with $\mathcal{O}(n^2+(n+1)m)$ arithmetic operations. It is much lower than that $\mathcal{O}(n^5m^{1.5})$ of the SDP method, where n and m are the dimension and sampling number, respectively.
3) We show that the proposed filter is equivalent to the famous SMF derived by Schweppe [8] when the nonlinear dynamic system is degenerated to linear. Besides, stability analysis of the proposed filter is derived for some special nonlinear dynamic systems.

Finally, two illustrative examples in the simulations show that the proposed DSMF can perform better.

The rest of this article is organized as follows. The problem of the SMF for the nonlinear systems is formulated in Section II. In Section III, a novel DSMF is developed for computing the state estimation ellipsoid. The semi-infinite programming in the DSMF is solved efficiently by the FW method in Section IV. Furthermore, we analyze the properties of the DSMF in Section V. Simulations and conclusions are given in Sections VI and VII, respectively. Some technical proofs are provided in the Appendixes.

Notation: For a square matrix $\mathbf{X}\succeq 0$ (respectively, $\mathbf{X}\succ 0$), $\mathbf{X}$ is semidefinite (respectively, positive definite). The superscript "T" denotes the transpose. "∂" is the gradient operator. $\mathrm{diag}(\cdot)$ and $\mathbf{I}$ stand for a block diagonal matrix and the identity matrix with an appropriate dimension, respectively. Ellipsoid is described as $\mathcal{E}=\{\mathbf{x}\in\mathcal{R}^n:(\mathbf{x}-\hat{\mathbf{x}})^T\mathbf{P}^{-1}(\mathbf{x}-\hat{\mathbf{x}})\le 1\}=\{\mathbf{x}:\mathbf{x}=\hat{\mathbf{x}}+\mathbf{E}\mathbf{u},\mathbf{P}=\mathbf{E}\mathbf{E}^T,\|\mathbf{u}\|\le 1\}$, where $\mathbf{P}$ and $\hat{\mathbf{x}}$ are the shape matrix and the center of the ellipsoid $\mathcal{E}$, respectively. The "size" of the ellipsoid is the function of the shape matrix $\mathbf{P}$, and it is denoted by $r(\mathbf{P})$. In this article, $r(\mathbf{P})$ is either $\log\det(\mathbf{P})$, which corresponds to the volume of the ellipsoid $\mathcal{E}$, or $\mathrm{tr}(\mathbf{P})$, which means the sum of squares of the semiaxis lengths of the ellipsoid $\mathcal{E}$.

II. Problem Formulation

In target tracking, one of the major objectives is to estimate the state trajectory of a target. In general, the dynamic model for target tracking describes the evolution of the target vector with respect to time. The most common state-space models with additive noises are given as follows:

$$\mathbf{x}_{k+1}=f_k(\mathbf{x}_k)+\mathbf{w}_k \tag{1}$$

$$\mathbf{y}_k=h_k(\mathbf{x}_k)+\mathbf{v}_k \tag{2}$$

where $\mathbf{x}_k\in\mathbb{R}^n$ and $\mathbf{y}_k\in\mathbb{R}^l$ are the target state and observation at time k, respectively. $f_k(\mathbf{x}_k)$ and $h_k(\mathbf{x}_k)$ are the nonlinear continuously differentiable functions of the state $\mathbf{x}_k$. Here, $\mathbf{w}_k$ and $\mathbf{v}_k$ are the process and observation noises, respectively.

The uncertain process noise $\mathbf{w}_k$ and measurement noise $\mathbf{v}_k$ are assumed to be bounded by the following ellipsoids:

$$\mathcal{W}_k=\{\mathbf{w}_k:\mathbf{w}_k^T\mathbf{Q}_k^{-1}\mathbf{w}_k\le 1\} \tag{3}$$

$$\mathcal{V}_k=\{\mathbf{v}_k:\mathbf{v}_k^T\mathbf{R}_k^{-1}\mathbf{v}_k\le 1\} \tag{4}$$

where $\mathbf{Q}_k$ and $\mathbf{R}_k$ are the shape matrix of the ellipsoids $\mathcal{W}_k$ and $\mathcal{V}_k$, respectively. Both of them are known symmetric positive-definite matrices.

Suppose that the initial state $\mathbf{x}_0$ belongs to a given bounding ellipsoid

$$\mathcal{E}_0=\{\mathbf{x}\in R^n:(\mathbf{x}-\hat{\mathbf{x}}_0)^T\mathbf{P}_0^{-1}(\mathbf{x}-\hat{\mathbf{x}}_0)\le 1\} \tag{5}$$

where $\hat{\mathbf{x}}_0$ is the center of the ellipsoid $\mathcal{E}_0$; $\mathbf{P}_0$ is the shape matrix of the ellipsoid $\mathcal{E}_0$ that is a known symmetric positive-definite matrix.

The goal of this article is to obtain an ellipsoid $\mathcal{E}_{k+1}$ to bound the state $\mathbf{x}_{k+1}$ at time $k+1$ by a recursion method. Specifically, at time k, assume that the SMF has obtained an ellipsoid $\mathcal{E}_k$ contains the state $\mathbf{x}_k$, i.e.,

$$\begin{aligned}\mathcal{E}_k&=\{\mathbf{x}:(\mathbf{x}-\hat{\mathbf{x}}_k)^T\mathbf{P}_k^{-1}(\mathbf{x}-\hat{\mathbf{x}}_k)\le 1\}\\&=\{\mathbf{x}:\mathbf{x}=\hat{\mathbf{x}}_k+\mathbf{E}_k\eta_k,\mathbf{P}_k=\mathbf{E}_k\mathbf{E}_k^T,\parallel\eta_k\parallel\le 1\}\end{aligned} \tag{6}$$

where $\hat{\mathbf{x}}_k$ is the center of the ellipsoid $\mathcal{E}_k$; $\mathbf{P}_k$ is a known symmetric positive-definite matrix. Based on the ellipsoid $\mathcal{E}_k$ and nonlinear state function, a predicted ellipsoid $\mathcal{E}_{k+1|k}$ in the prediction step can be derived, which is

$$\begin{aligned}\mathcal{E}_{k+1|k}&=\{\mathbf{x}:(\mathbf{x}-\hat{\mathbf{x}}_{k+1|k})^T\mathbf{P}_{k+1|k}^{-1}(\mathbf{x}-\hat{\mathbf{x}}_{k+1|k})\le 1\}\\&=\{\mathbf{x}:\mathbf{x}=\hat{\mathbf{x}}_{k+1|k}+\mathbf{E}_{k+1|k}\eta_{k+1|k},\\&\qquad\mathbf{P}_{k+1|k}=\mathbf{E}_{k+1|k}\mathbf{E}_{k+1|k}^T,\parallel\eta_{k+1|k}\parallel\le 1\}\end{aligned} \tag{7}$$

where $\hat{\mathbf{x}}_{k+1|k}$ is the center of the ellipsoid $\mathcal{E}_{k+1|k}$ and $\mathbf{P}_{k+1|k}$ is a symmetric positive-definite matrix. Then, we use the nonlinear measurement function (2) and measurement $\mathbf{y}_k$ to obtain the

updated ellipsoid $\mathcal{E}_{k+1}$ at time $k+1$ in the measurement update step, which is defined as follows:

$$\begin{aligned}\mathcal{E}_{k+1} &= \{\mathbf{x} : (\mathbf{x}-\hat{\mathbf{x}}_{k+1})^T \mathbf{P}_{k+1}^{-1}(\mathbf{x}-\hat{\mathbf{x}}_{k+1}) \le 1\} \\ &= \{\mathbf{x} : \mathbf{x} = \hat{\mathbf{x}}_{k+1} + \mathbf{E}_{k+1}\eta_{k+1}, \\ &\quad \mathbf{P}_{k+1} = \mathbf{E}_{k+1}\mathbf{E}_{k+1}^T, \| \eta_{k+1} \| \le 1\} \end{aligned} \tag{8}$$

where $\hat{\mathbf{x}}_{k+1}$ is the center of the ellipsoid $\mathcal{E}_{k+1}$ and $\mathbf{P}_{k+1}$ is a symmetric positive-definite matrix.

In order to obtain a tighter ellipsoid, in the next section, a novel SMF is developed by using an advanced optimization method to deal with the nonlinear function, rather than linearizing it.

III. DSMF Without Linearization

In this section, a DSMF is proposed for the nonlinear system, which does not need to linearize the nonlinear function, and it derives a predicted ellipsoid and an updated ellipsoid by solving the semi-infinite programming, respectively.

A. Prediction Step

Now, we consider the prediction step in the DSMF; specifically, based on the ellipsoid $\mathcal{E}_k$ and the state equation at time k, we determine a predicted ellipsoid $\mathcal{E}_{k+1|k}$ that covers the set of state at time $k+1$. In general, there exist many ellipsoids containing the reachable set of states; however, finding a tighter predicted ellipsoid is difficult, especially in the nonlinear system.

The traditional ESMF [20] is linearizing the function f_k in (1) about the current state estimate $\hat{\mathbf{x}}_k$ [defined in (6)] yields

$$\begin{aligned}\mathbf{x}_{k+1} = f_k(\hat{\mathbf{x}}_k) + \frac{\partial f_k(\mathbf{x}_k)}{\partial \mathbf{x}}\Big|_{\mathbf{x}_k=\hat{\mathbf{x}}_k}(\mathbf{x}_k-\hat{\mathbf{x}}_k) \\ + R_k^f(\mathbf{x}_k,\hat{\mathbf{x}}_k) + \mathbf{w}_k \end{aligned} \tag{9}$$

where $R_k^f(\mathbf{x}_k,\hat{\mathbf{x}}_k)$ is the high-order Lagrange remainder [20] and is written as

$$\begin{aligned}R_k^f(\mathbf{x}_k,\hat{\mathbf{x}}_k) = (\mathbf{x}_k-\hat{\mathbf{x}}_k)^T \frac{1}{2}\frac{\partial^2 f_k(\bar{X}_k)}{\partial \mathbf{x}}(\mathbf{x}_k-\hat{\mathbf{x}}_k) \\ \forall\ \bar{X}_k = \hat{\mathbf{x}}_k + \theta_k(\mathbf{x}_k-\hat{\mathbf{x}}_k), 0 \le \theta_k \le 1. \end{aligned} \tag{10}$$

The other form for the remainder [28] is

$$R_k^f(\mathbf{x}_k,\hat{\mathbf{x}}_k) = f_k(\mathbf{x}_k) - f_k(\hat{\mathbf{x}}_k) - \frac{\partial f_k(\mathbf{x}_k)}{\partial \mathbf{x}}\Big|_{\mathbf{x}_k=\hat{\mathbf{x}}_k}(\mathbf{x}_k-\hat{\mathbf{x}}_k). \tag{11}$$

In [20], the higher order term $R_k^f(\mathbf{x}_k,\hat{\mathbf{x}}_k)$ in (10) can be bounded by an interval, and the interval is further bounded by an ellipsoid. However, this method is conservative, and sometimes, it is difficult to calculate the Hessian matrix. For the second form in (11), it is hard to analyze its property and find a tighter ellipsoid to contain the remainder.

In this section, an optimization technique is used to avoid the above difficulty. Note that if the nonlinear state transform function f_k is uniformly continuous, then $\mathcal{F}_k = \{f_k(\mathbf{x}_k) : \mathbf{x}_k \in \mathcal{E}_k\}$ is a compact set. Compared with linearizing the nonlinear function f_k, an ellipsoid $\mathcal{E}_{f_k}$ containing the nonlinear transformation $\mathcal{F}_k$ can be directly derived, i.e., $\mathcal{E}_{f_k} \supseteq \mathcal{F}_k$, by solving the following optimization problem:

$$\min\ f(\mathbf{P}_{f_k}) \tag{12}$$

$$\text{s.t. } [\mathbf{x}_{f_k}-\hat{\mathbf{x}}_{f_k}]^T \mathbf{P}_{f_k}^{-1}[\mathbf{x}_{f_k}-\hat{\mathbf{x}}_{f_k}] \le 1$$

$$\forall \mathbf{x}_{f_k} \in \mathcal{F}_k \tag{13}$$

with variables $\hat{\mathbf{x}}_{f_k}$ and $\mathbf{P}_{f_k}$, where $\hat{\mathbf{x}}_{f_k}$ and $\mathbf{P}_{f_k}$ are the center and shape matrix of the ellipsoid $\mathcal{E}_{f_k}$, respectively. Fig. 1 gives us an illustration of the optimization problem (12), (13). The optimization problem has the property that there are two variables appearing in infinitely many constraints; then, it is called the semi-infinite programming. In the next section, we will provide some methods to solve the optimization problem (12), (13) and obtain the ellipsoid $\mathcal{E}_{f_k}$. Thus, it is assumed that the ellipsoid $\mathcal{E}_{f_k}$ is obtained in this subsection.

As well known, the other form of the ellipsoid $\mathcal{E}_{f_k}$ is

$$\mathcal{E}_{f_k} = \{\mathbf{x} : \mathbf{x} = \hat{\mathbf{x}}_{f_k} + \mathbf{E}_{f_k}\eta_{f_k}, \|\eta_{f_k}\| \le 1, \mathbf{E}_{f_k}\mathbf{E}_{f_k}^T = \mathbf{P}_{f_k}\}.$$

Then, there exists η_{f_k} such that $f_k(\mathbf{x}_k) = \hat{\mathbf{x}}_{f_k} + \mathbf{E}_{f_k}\eta_{f_k}$ by $f_k(\mathbf{x}_k) \in \mathcal{E}_{f_k}$. Therefore, the nonlinear equation (1) can be translated into a linear equation as follows:

$$\mathbf{x}_{k+1} = \hat{\mathbf{x}}_{f_k} + \mathbf{E}_{f_k}\eta_{f_k} + \mathbf{w}_k. \tag{14}$$

According to (14), it can be asserted that

$$\mathbf{x}_{k+1} \in \mathcal{E}_{f_k} \oplus \mathcal{W}_k. \tag{15}$$

Although the Minkowski sum $\mathcal{E}_{f_k} \oplus \mathcal{W}_k$ is not an ellipsoid, in [9], Kurzhanski and Vâlyi have shown that there exists the ellipsoid $\mathcal{E}_{k+1|k}$ that is an external approximation of the Minkowski sum, i.e.,

$$\mathcal{E}_{f_k} \oplus \mathcal{W}_k \subseteq \mathcal{E}_{k+1|k} \tag{16}$$

where the center and shape matrix of the predicted ellipsoid $\mathcal{E}_{k+1|k}$ are calculated as follows:

$$\hat{\mathbf{x}}_{k+1|k} = \hat{\mathbf{x}}_{f_k} \tag{17}$$

$$\mathbf{P}_{k+1|k}(p_k) = (1+p_k^{-1})\mathbf{P}_{f_k} + (1+p_k)\mathbf{Q}_k \tag{18}$$

for any $p_k > 0$, respectively. Now, we may select an optimal external ellipsoid $\mathcal{E}_{k+1|k}$ relative to some criteria by solving

$$\min_{p_k>0}\ r\left(\mathbf{P}_{k+1|k}(p_k)\right). \tag{19}$$

Fortunately, there exists a unique ellipsoid with minimal sum of squares of semiaxes containing the Minkowski sum [14], where the optimal value of p_k is defined by p_k^*, and

$$p_k^* = \frac{\sqrt{\text{tr}(\mathbf{P}_{f_k})}}{\sqrt{\text{tr}(\mathbf{Q}_k)}}. \tag{20}$$

Let $\mathbf{P}_{k+1|k}$ denote as $\mathbf{P}_{k+1|k}(p_k^*)$; hence, the optimal predicted ellipsoid $\mathcal{E}_{k+1|k}$ has been found.

B. Measurement Update Step

In the measurement update step, the goal is to determine a minimal updated ellipsoid $\mathcal{E}_{k+1}$ that contains the states, which is consistent with both the prediction $\mathcal{E}_{k+1|k}$ and the measurement at time $k+1$.

If we linearize the nonlinear measurement function h_{k+1} about the current state estimate $\hat{\mathbf{x}}_{k+1|k}$, it yields

$$\mathbf{y}_{k+1} = h_{k+1}(\hat{\mathbf{x}}_{k+1|k}) + R^h_{k+1}(\mathbf{x}_{k+1}, \hat{\mathbf{x}}_{k+1|k}) + \mathbf{w}_{k+1} + \frac{\partial h_{k+1}(\mathbf{x}_{k+1})}{\partial \mathbf{x}}\Big|_{\mathbf{x}_{k+1}=\hat{\mathbf{x}}_{k+1|k}} (\mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k}) \quad (21)$$

where $R^h_{k+1}(\mathbf{x}_{k+1}, \hat{\mathbf{x}}_{k+1|k})$ is the high-order remainder, and

$$R^h_{k+1}(\mathbf{x}_{k+1}, \hat{\mathbf{x}}_{k+1|k}) = \eta^T_{k+1|k}\mathbf{E}^T_{k+1|k}\frac{1}{2}\frac{\partial^2 f(\bar{X}_{k+1|k})}{\partial \mathbf{x}}\mathbf{E}_{k+1|k}\eta_{k+1|k}. \quad (22)$$

Since the larger predicted ellipsoid means that the bigger shape matrix $\mathbf{E}_{k+1|k}$, from (22), it shows that the remainder becomes larger with the larger predicted ellipsoid, which needs to find a larger measurement ellipsoid covering the remainder. Finally, the intersection of the measurement ellipsoid and the predicted ellipsoid is larger, which brings a worse updated ellipsoid to bound the intersection. To obtain a better result, we can solve a semi-infinite optimization problem to cover the nonlinear transformation of predicted ellipsoid directly, rather than linearizing the nonlinear measurement function.

For the convenience of analysis, assume that the nonlinear function h_{k+1} exists a continuous inverse function h^{-1}_{k+1}, and this assumption can be relaxed for the special and important field, mobile robot location, in Section VI. Then, (2) can be rewritten as follows:

$$h^{-1}_{k+1}(\mathbf{y}_{k+1} - \mathbf{v}_{k+1}) = \mathbf{P}_{\mathbf{x}}\mathbf{x}_{k+1} \quad (23)$$

where $\mathbf{P}_{\mathbf{x}}$ is the projection matrix. For example, when the measurement function [29] is denoted as

$$h(\mathbf{x}) = \begin{bmatrix} r \\ \theta \end{bmatrix} = \begin{bmatrix} \sqrt{(x-a)^2 + (y-b)^2} \\ \arctan\frac{y-b}{x-a} \end{bmatrix} \quad (24)$$

where $\mathbf{x} = [x, \dot{x}, y, \dot{y}, \ddot{x}, \ddot{y}]$, (x, y) is the target position, and (a, b) is the sensor position. Then

$$h^{-1} = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r\cos(\theta) + a \\ r\sin(\theta) + b \end{bmatrix} = \mathbf{P}_{\mathbf{x}}\mathbf{x} \quad (25)$$

where

$$\mathbf{P}_{\mathbf{x}} = \begin{bmatrix} 1\,0\,0\,0\,0\,0 \\ 0\,0\,1\,0\,0\,0 \end{bmatrix}.$$

Based on the uncertain set $\mathbf{v}_{k+1}$ of the measurement noise, we hope to find a minimum measurement ellipsoid $\mathcal{E}(\hat{\mathbf{z}}_{k+1}, \mathbf{P}_{z_{k+1}})$ to contain the left-hand side of (23), which can be described as follows:

$$\min\ r(\mathbf{P}_{z_{k+1}}) \quad (26)$$

$$\text{s.t. } \mathcal{C}_{k+1} \subseteq \mathcal{E}(\hat{\mathbf{z}}_{k+1}, \mathbf{P}_{z_{k+1}}) \quad (27)$$

where $\mathcal{C}_{k+1} = \{h^{-1}_{k+1}(\mathbf{y}_{k+1} - \mathbf{v}_{k+1}) : \mathbf{v}_{k+1} \in \mathcal{V}_{k+1}\}$. $f(\mathbf{P}_{z_{k+1}})$ is the objective function, which is the "size" of the shape matrix $\mathbf{P}_{z_{k+1}}$.

The optimization problem (26), (27) is equivalent to

$$\min\ r(\mathbf{P}_{z_{k+1}}) \quad (28)$$

$$\text{s.t. } (\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1})^T\mathbf{P}^{-1}_{z_{k+1}}(\mathbf{z}_{k+1} - \hat{\mathbf{z}}_{k+1}) \le 1$$

$$\forall\ \mathbf{z}_{k+1} \in \mathcal{C}_{k+1} \quad (29)$$

which is also a semi-infinite optimization problem, and it has the similar form with the optimization problem (12), (13). Then, we can solve it at next section.

Since $\mathbf{P}_{\mathbf{x}}\mathbf{x}_{k+1}$ belongs to $\mathcal{E}(\hat{\mathbf{z}}_{k+1}, \mathbf{P}_{z_{k+1}})$ by (23), and it is equivalent to

$$-\hat{\mathbf{z}}_{k+1} = -\mathbf{P}_{\mathbf{x}}\mathbf{x}_{k+1} + \eta^z_{k+1}, \eta^{z^T}_{k+1}\mathbf{P}^{-1}_{z_{k+1}}\eta^z_{k+1} \le 1. \quad (30)$$

The nonlinear measurement equation (2) can be approximated by the above linear equation.

The measurement update step is to find an updated ellipsoid $\mathcal{E}(\hat{\mathbf{x}}_{k+1}, \mathbf{P}_{k+1})$ containing the intersection of predicted ellipsoid $\mathcal{E}_{k+1|k}$ and measurement ellipsoid $\mathcal{E}(\hat{\mathbf{z}}_{k+1}, \mathbf{P}_{z_{k+1}})$, where $\hat{\mathbf{x}}_{k+1}$ and $\mathbf{P}_{k+1}$ are the center and shape matrix of the updated ellipsoid.

According to Appendix A, the formula of the measurement update step can be written as follows:

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_{k+1|k} + \frac{\mathbf{P}_{k+1|k}}{1-\rho_{k+1}}\mathbf{P}^T_{\mathbf{x}} \cdot \left[\mathbf{P}_{\mathbf{x}}\frac{\mathbf{P}_{k+1|k}}{1-\rho_{k+1}}\mathbf{P}^T_{\mathbf{x}} + \frac{\mathbf{P}_{z_{k+1}}}{\rho_{k+1}}\right]^{-1} \cdot (\hat{\mathbf{z}}_{k+1} - \mathbf{P}_{\mathbf{x}}\hat{\mathbf{x}}_{k+1|k}), \quad (31)$$

$$\bar{\mathbf{P}}_{k+1} = \left[(1-\rho_{k+1})\mathbf{P}^{-1}_{k+1|k} + \rho_{k+1}\mathbf{P}^T_{\mathbf{x}}\mathbf{P}^{-1}_{z_{k+1}}\mathbf{P}_{\mathbf{x}}\right]^{-1}, \quad (32)$$

$$\delta_{k+1} = (\hat{\mathbf{z}}_{k+1} - \mathbf{P}_{\mathbf{x}}\hat{\mathbf{x}}_{k+1|k})^T \cdot \left[\mathbf{P}_{\mathbf{x}}\frac{\mathbf{P}_{k+1|k}}{1-\rho_{k+1}}\mathbf{P}^T_{\mathbf{x}} + \frac{\mathbf{P}_{z_{k+1}}}{\rho_{k+1}}\right]^{-1} \cdot (\hat{\mathbf{z}}_{k+1} - \mathbf{P}_{\mathbf{x}}\hat{\mathbf{x}}_{k+1|k}), \quad (33)$$

$$\mathbf{P}_{k+1} = (1-\delta_{k+1})\bar{\mathbf{P}}_{k+1}, 0 \le \rho_{k+1} \le 1. \quad (34)$$

The optimal parameter ρ^*_{k+1} can be obtained by solving the following problem:

$$\min\ r(\mathbf{P}_{k+1}) \quad (35)$$

$$\text{s.t. } 1 \ge \rho_{k+1} \ge 0. \quad (36)$$

Substituting ρ^*_{k+1} into (31)–(34), the final center and the shape matrix of the updated ellipsoid can be obtained.

Remark 1: Searching for an outer-bounding ellipsoid $\mathcal{E}_{k+1}$ is reduced to a 1-D optimization by (35) and (36). Meanwhile, the optimization problem (35), (36) is a convex optimization problem, which can be solved efficiently by many methods, such as the golden section method.

In this example, we focus on the updated ellipsoid in the measurement update step for different nonlinear SMF methods.

Example 1: The center and shape matrix of the predicted ellipsoid are $[1020]^T$ and $\sigma\mathbf{I}$, respectively. Here, σ is a parameter. The nonlinear measurement function in (2) is

$$h_k(\mathbf{x}) = \begin{bmatrix} \sqrt{\mathbf{x}_1^2 + \mathbf{x}_2^2} \\ \arctan\frac{\mathbf{x}_2}{\mathbf{x}_1} \end{bmatrix}$$

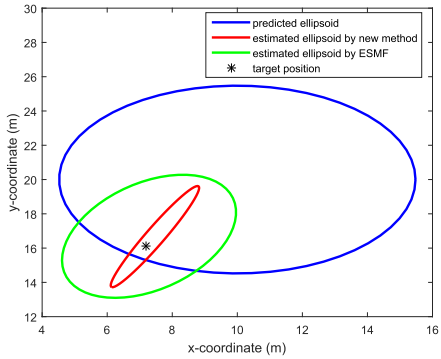

Fig. 2. Estimated ellipsoid with the proposed method and the ESMF with $\sigma = 30$.

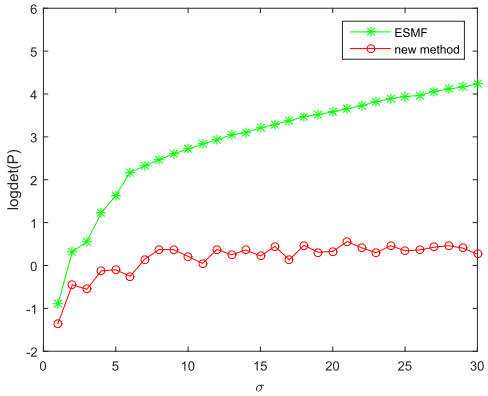

Fig. 3. Size (logdet) of the estimated ellipsoid is plotted as a function of σ for 50 Monte Carlo runs.

where $\mathbf{x} = [x_1, x_2]^T$ is the true state, which is sampled from the predicted ellipsoid, and $\mathbf{v}$ is the measurement noise with shape matrix $\mathbf{R} = \mathrm{diag}(10, 1)$. Generate a measurement y from the measurement equation (2). According to the predicted ellipsoid and measurement, the ESMF and the new method are exploited to get the updated ellipsoid $\mathcal{E}$ with shape matrix $\mathbf{P}$.

1) The new method can obtain a smaller updated ellipsoid in Fig. 2.
2) From Fig. 3, we can see that with the increasing of σ, which means that the predicted ellipsoid becomes larger, the size of the estimated ellipsoid increases. However, the new method gets a stable estimated ellipsoid.

The reason may be that our method does not need to linearize the nonlinear measurement function, which does not bring much uncertainty.

Note that if the intersection of the predicted ellipsoid $\mathcal{E}_{k+1|k}$ and the measurement ellipsoid $\mathcal{E}(\hat{\mathbf{z}}_{k+1}, \mathbf{P}_{z_{k+1}})$ is empty, it means that the measurement $\mathbf{y}_k$ is not compatible with the dynamic system (1), (2). This situation may happen when the true process noise and measurement noise outstrip the assumed bounds. In this case, we use the predicted ellipsoid $\mathcal{E}_{k+1|k}$ instead of the updated ellipsoid $\mathcal{E}_{k+1}$ in the next step.

A summary of the proposed DSMF is in Algorithm 1, and we can apply it to target tracking. In Section IV, we provide an efficient method to solve the semi-infinite optimization problem in Algorithm 1.

Algorithm 1: Dual Set Membership Filter.

Input: The nonlinear state function f_k, measurement function h_k, the measurement $\mathbf{y}_k$, the center $\mathbf{x}_0$ and shape matrix $\mathbf{P}_0$ of initial ellipsoid;

Output:

1: **for** each $k = 1 : K$ **do**
2: Solve optimization problem (12), (13) to obtain $\hat{\mathbf{x}}_{f_k}$ and $\mathbf{P}_{f_k}$ (see Algorithm 3);
3: Calculate the predicted ellipsoid $\mathcal{E}_{k+1|k}$ by (17) and (18);
4: Solve optimization problem (28), (29) to obtain $\hat{\mathbf{z}}_k$ and $\mathbf{P}_{z_k}$ (see Algorithm 3);
5: Calculate the updated ellipsoid $\mathcal{E}_{k+1}$ by (31)–(34);
6: **end for**
7: **return** $\hat{\mathbf{x}}_{k+1}$ and $\mathbf{P}_{k+1}$;

IV. Solving the Semi-Infinite Programming

To overcome the difficulty caused by the nonlinear function, the nonlinear dynamic system has been translated into the linear system (14), (30) by solving the semi-infinite programming (12), (13) and (28), (29). In this section, we focus on discussing how to solve the semi-infinite programming.

First, the optimization problems (12), (13) and (28), (29) can be unified as

$$\min_{\mathbf{P}, \hat{\mathbf{y}}} \log\det(\mathbf{P}) \tag{37}$$

$$\text{s.t. } (\mathbf{y} - \hat{\mathbf{y}})^T \mathbf{P}^{-1} (\mathbf{y} - \hat{\mathbf{y}}) \leq n \forall \mathbf{y} \in \mathcal{C}. \tag{38}$$

Here, the vector $\hat{\mathbf{y}}$ and matrix $\mathbf{P}$ are the optimization variables, and they are also the center and the shape matrix of ellipsoid $\mathcal{E}$. n is the dimension of $\mathbf{y}$. The set $\mathcal{C}$ is defined as

$$\mathcal{C} = \{\mathbf{y} : \mathbf{y} = s(\mathbf{x}), \mathbf{x} \in \mathcal{E}_0\} \tag{39}$$

$$\mathcal{E}_0 = \{\mathbf{y} : (\mathbf{y} - \hat{\mathbf{y}}_0)^T \mathbf{P}_0^{-1} (\mathbf{y} - \hat{\mathbf{y}}_0) \leq 1\} \tag{40}$$

where s is an arbitrary nonlinear continuous function, and $\hat{\mathbf{y}}_0$ and $\mathbf{P}_0$ are the center and shape matrix of the ellipsoid $\mathcal{E}_0$, respectively. If the objective function (37) is the trace function, i.e., $\mathrm{tr}(\mathbf{P})$, the similar result and algorithm can be obtained.

A. First-Order Optimal Condition

In this subsection, we discuss the problem (37), (38) from a theoretical point of view, and its dual problem helps us to design a fast algorithm to solve the problem (37), (38). Thus, the proposed filter is called as the DSMF. The following proposition shows that strong duality holds.

Proposition 1: The optimal solutions of problem (37), (38) are

$$\hat{\mathbf{y}}^* = \int_{\mathcal{C}} \mathbf{y} d\mu^*, \mathbf{P}^* = \int_{\mathcal{C}} \mathbf{y}\mathbf{y}^T d\mu^* - \hat{\mathbf{y}}^* \hat{\mathbf{y}}^{*T} \tag{41}$$

where μ^* is a measure and the optimal solution of the following optimization problem:

$$\max_{\mu} \log\det\left(\int_{\mathcal{C}\times\{1\}} \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu\right) \tag{42}$$

$$\text{s.t.} \int_{\mathcal{C}\times\{1\}} d\mu = 1, \mu \geq 0 \tag{43}$$

where $\tilde{\mathbf{y}} = [\mathbf{y}^T, 1]^T$.

Proof: See Appendix B.

Remark 2: Compared the optimization problems (37) and (38) with (42) and (43), the constraint function in (43) is a linear function on a continuous measure μ, which is easier to find the feasible solution. If μ is a discrete measure, the constraint (43) is a simplex, which offers guidance on the efficient first-order algorithm.

Since the proof of Proposition 1 needs the following two lemmas, and these two lemmas are important to our algorithm, we first show them as follows.

Lemma 1: Let $d = n + 1$; the minimum of problem

$$\min_{\mathbf{H}} -\log\det(\mathbf{H}) \tag{44}$$

$$\text{s.t. } \tilde{\mathbf{y}}^T\mathbf{H}\tilde{\mathbf{y}} \leq d, \forall \tilde{\mathbf{y}} \in \Omega = \mathcal{C} \times \{1\} \tag{45}$$

is obtained at $\mathbf{H}^*$ if and only if $\mathbf{H}^*$ is feasible, and there exists a μ^* such that

$$\mathbf{H}^{*-1} = \int_{\Omega} \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu^* \tag{46}$$

$$\int_{\Omega} (\tilde{\mathbf{y}}^T\mathbf{H}^*\tilde{\mathbf{y}} - d) d\mu^* = 0 \tag{47}$$

$$\mu^* \geq 0. \tag{48}$$

Proof: See Appendix B.

Lemma 2: The dual problem corresponding to the primal problem (44), (45) is

$$\max_{\mu} \log\det\left(\int_{\Omega} \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu\right) \tag{49}$$

$$\text{s.t.} \int_{\Omega} \mu = 1, \mu \geq 0. \tag{50}$$

Proof: See Appendix B.

Indeed, it suffices to consider the discrete measures μ from John's optimality conditions [27], which puts positive measure on a finite points of the set Ω. Next, some points $\tilde{\mathbf{y}}_i$ are randomly sampled from the set Ω so that the optimization problem (42), (43) can be approximated by

$$\max_{\mu_i} \log\det\left(\sum_{i=1}^{m} \mu_i \tilde{\mathbf{y}}_i \tilde{\mathbf{y}}_i^T\right) \tag{51}$$

$$\text{s.t.} \sum_{i=1}^{m} \mu_i = 1, \mu_i \geq 0. \tag{52}$$

Since $\Omega = \mathcal{C} \times \{1\}$ and $\tilde{\mathbf{y}} = [\mathbf{y}^T, 1]^T$, if we want to get points $\tilde{\mathbf{y}}_i$ from the set Ω, it is sufficient to get points $\mathbf{y}_i$ from the set $\mathcal{C}$, and $i = 1, \ldots, m$.

Using statistical learning techniques, in [30], Calafiore and Campi provide an explicit bound on the probability of the set of original constraints that are possibly violated by the randomized solution. When the number of samples m satisfies $m \geq \frac{n}{\eta\beta} - 1$, with probability no smaller than $1 - \beta$, the approximated problem (51), (52) returns an optimal solution, which is η-level robustly feasible. Here, n is the dimension of the decision variable. Therefore, the solution of the optimization problem (49), (50) can be made approximately feasible for the semi-infinite optimization problem (51), (52) by sampling a sufficient number of constraints.

B. Relationship Between the DSMF and the UKF

In this subsection, we first consider three cases for obtaining the points $\mathbf{y}_i$ from the set $\mathcal{C}$.

1) For a general function s defined in (39), it is easy to sample the interior points or boundary points $\mathbf{x}_i$ from the ellipsoid $\mathcal{E}_0$; then, the points in the set $\mathcal{C}$ can be directly obtained by $\mathbf{y}_i = s(\mathbf{x}_i)$, and $i = 1, \ldots, m$.
2) For some particular nonlinear functions, such as the homeomorphic function [31], it takes the boundary of the ellipsoid $\mathcal{E}_0$ to the boundary of the set $\mathcal{C}$. Since $\int_{\Omega}(\tilde{\mathbf{y}}^T\mathbf{H}^*\tilde{\mathbf{y}} - d)d\mu^* = 0$ in (47), when $\tilde{\mathbf{y}}^T\mathbf{H}^*\tilde{\mathbf{y}} = d$, $\mu \neq 0$; otherwise, $\mu = 0$. To obtain the points that making $\mu \neq 0$, it is sufficient for getting the boundary points of the set $\mathcal{C}$.
3) For some common nonlinear functions in 2-D or 3-D radar systems [28], it is only required to take samples on the boundary (zero measure) of the ellipsoid $\mathcal{E}_0$ based on (47) in Lemma 1; then, it is enough to use them to get the boundary points $\mathbf{y}_i$ in $\mathcal{C}$ and $i = 1, \ldots, m$.

Next, we show the relationship between the DSMF and the UKF. As we know, for the UKF, a set of points (sigma points) are chosen, and the nonlinear function is applied to these points, in turn, to yield a cloud of transformed points. Then, the statistics of the transformed points can be approximated to form an estimate of the nonlinearly transformed mean and covariance [5].

For our methods, a set of interior points or boundary points are chosen from the ellipsoid $\mathcal{E}_0$, and the nonlinear function is also applied to them; however, we want to find an ellipsoid to cover all these nonlinear transformed points, which is formulated as the optimization problem (51), (52). Thus, compared with the UKF, the DSMF can obtain a robust estimate of the true state by solving the optimization problem (51), (52).

C. FW Method

The FW or conditional gradient algorithm is a projection-free method; it solves the problem (51), (52) without requiring to compute a projection onto the feasible set.

Let $g(\mu) = \log\det(\sum_{i=1}^{m} \mu_i \tilde{\mathbf{y}}_i \tilde{\mathbf{y}}_i^T)$ and $M = \{\mu : \sum_{i=1}^{m} \mu_i = 1, \mu_i \geq 0\}$. If we make a first-order linear (Taylor) approximation to the objective function $g(\mu)$ at the current solution μ^t, then

$$g(\mu) \approx g(\mu^t) + \omega^{t^T}(\mu - \mu^t) \tag{53}$$

where $\omega^t := \omega(\mu^t) = \partial g(\mu^t)$. When we maximize this linear function over the unit simplex M, the analysis solution is obtained at a vertex, and the analysis solution is at a vertex, i.e., unit vector e_i, which means the ith component is 1, the other is 0. Hence, we might wish to move along the line from μ^t toward e_i

Algorithm 2 FW Algorithm

Input: Let $\mu^0 \in M$

Output:

1: **for** each $t = 1 : T$ **do**
2: Compute $e_i^t : \arg\min_{s \in M} \langle s, \partial g(\mu^t) \rangle$;
3: Let $d_t = e_i^t - \mu^t$;
4: Obtain the optimal step $\gamma_t = \arg\min_{\gamma \in [0,1]} g(\mu^t + \gamma d_t)$;
5: Update $\mu^{t+1} = \mu^t + \gamma_t d_t$.
6: **end for**
7: **return** μ^T;

Algorithm 3 First-Order Method for Semi-Infinite Programming

Input: The number of samples m; The set $\mathcal{C}$;

Output: Obtain ellipsoid $\mathcal{E}$ to contain set $\mathcal{C}$;

1: Generate samples $\mathbf{y}_1, \ldots, \mathbf{y}_m$ from set $\mathcal{C}$, and let $\tilde{\mathbf{y}}_i = [\mathbf{y}_i^T, 1]^T$;
2: Solve the optimization problem (51), (52) to obtain μ_i^* by the FW algorithm;
3: **return** $\hat{\mathbf{y}} = \sum_{i=1}^m \mu_i^* \mathbf{y}_i, \mathbf{P} = \sum_{i=1}^m \mu_i^* \mathbf{y}_i \mathbf{y}_i^T - \hat{\mathbf{y}}\hat{\mathbf{y}}^T$;

to update μ^{t+1}. The pseudocode for the FW algorithm is shown in Algorithm 2.

In [26], Ahipasaoglu *et al.* have shown the local linear convergence of the FW algorithm for the optimization problem (51), (52), and it has a decisive advantage in a large-scale problem. Based on Proposition 1, Algorithm 3 can be used to solve the semi-infinite programming problems (12), (13) and (28), (29).

D. FW Method Versus SDP

According to the Schur complement, the original optimization problem (37), (38) by introducing variables $\mathbf{P} = \mathbf{B}^{-1}$ and $\hat{\mathbf{y}} = \mathbf{P}\hat{\mathbf{b}}$ can be equivalent to

$$\min_{\mathbf{B}, \hat{\mathbf{b}}} -\log\det(\mathbf{B}) \tag{54}$$

$$\text{s.t.} \begin{bmatrix} n & \mathbf{y}^T\mathbf{B} - \hat{\mathbf{b}}^T \\ \mathbf{B}\mathbf{y} - \hat{\mathbf{b}} & \mathbf{B} \end{bmatrix} \geq 0 \quad \forall\, \mathbf{y} \in \mathcal{C} \tag{55}$$

where $\mathbf{B}$ and $\hat{\mathbf{b}}$ are the optimal variable. In order to solve the problem (54), (55), we can take samples from the boundary and the interior of the set $\mathcal{C}$, so that a finite set of $\mathbf{y}_1, \ldots, \mathbf{y}_m$ can be obtained; then, the infinite constraint (55) can be approximated by m constraints based on $\mathbf{y}_1, \ldots, \mathbf{y}_m$. Specifically, the optimization problem (54), (55) can be approximated by an SDP problem

$$\min_{\mathbf{B}, \hat{\mathbf{b}}} -\log\det(\mathbf{B}) \tag{56}$$

$$\text{s.t.} \begin{bmatrix} n & \mathbf{y}_i^T\mathbf{B} - \hat{\mathbf{b}}^T \\ \mathbf{B}\mathbf{y}_i - \hat{\mathbf{b}} & \mathbf{B} \end{bmatrix} \geq 0$$

$$\forall\, i = 1, \ldots, m. \tag{57}$$

TABLE I
MEAN SOLUTION RUNNING TIME OF SDPT3 AND FW

n	m	SDP	FW
2	50	0.7236	0.0027
2	100	1.2269	0.0037
2	200	2.2153	0.0076
2	400	4.1819	0.0048
2	600	6.2677	0.0115
2	800	8.4777	0.0141
2	1000	10.7986	0.0070
6	50	0.8008	0.0108
6	100	1.3663	0.0119
6	200	2.5366	0.0185
6	400	5.0538	0.0260
6	600	7.7717	0.0305
6	800	10.7293	0.0331
6	1000	13.9558	0.0416

In this problem, the variable is the matrix $\mathbf{B}$ and $\hat{\mathbf{b}}$; thus, the dimension of the optimization variable is $N = n(n+1)/2 + n$.

When the inequalities (57) are combined into one large inequalities, the dimension of the constraint matrix is $m(n+1)$. Boyd *et al.* [32] have developed a (primal-dual) interior-point method that solves (56) and (57) by exploiting the problem structure. They prove the worst-case estimate of $\mathcal{O}(N^{2.75}m^{1.5})$ arithmetic operations to solve the problem to a given accuracy. However, each iteration in the FW method requires $\mathcal{O}(n^2 + (n+1)m)$ arithmetic operations [27], where n and m are the dimension and the sampling number, respectively. The cheapness of the iterations in the FW method, together with its (relatively) attractive local linear convergence properties [26], [27], leads to that it is efficiency for large number m.

Example 2: There are m random vectors $\mathbf{y}_i, i = 1, \ldots, m$, that are generated by the standard uniform distribution. Assume that the dimension of $\mathbf{y}_i$ is n. We solve the SDP problem (56), (57) and the optimization problem (51), (52) by the SDPT3 algorithm using the CVX platform and the FW algorithm, respectively. The results presented in Table I are mean solution times of 20 Monte Carlo runs by the SDP solver in the third column and by the FW algorithm in the fourth column. It is shown that the first-order FW algorithm dominates the SDP method; sometimes, it is more than 200 times faster. Thanks to the faster FW algorithm, the proposed DSMF can work online.

The solving of the semi-infinite programming involves two approximations: one is by sampling a list of points, and the other is by the first-order approximation. The following remark is to clarify the sacrifice of accuracy as well as the compromise with the time complexity.

Remark 3:

1) First, it is a time-consuming and challenging problem to find the optimal solution of the semi-infinite programming (49), (50) [33]; then, we have to approximate it by discretization methods [30], [34], which randomly sample a sufficient number of constraints from the infinite constraints. In addition, Calafiore and Campi [30] have provided an explicit bound on the measure of the set of original constraints that are possibly violated by the randomized solution, and they prove that this measure rapidly decreases to zero as the number of samples m is

increased. Moreover, for some common nonlinear functions in a 2-D or 3-D radar system, it is only required to take samples on the boundary (zero measure) of the set $\mathcal{C}$. Then, it can reduce a lot of redundant samples so that the computational time can be decreased much more.

2) Second, the reason for using the first-order method is that the second-order method has higher time complexity to compute the Hessian matrix, which limits its online usage for large-scale problems (see Table I). To reduce the time complexity, we propose the first-order FW method to solve the optimization problem (51), (52), which only needs to use the first-order approximation without computing the Hessian matrix; thus, it has a lower computational complexity. Note that, in [27], Todd has proved that the FW algorithm with the first-order approximation enjoys global convergence. Thus, this approximation does not scarify the accuracy from the theory to obtain an ϵ-optimal solution.

In summary, the more samples and the smaller ϵ we use, the higher accuracy for the solution of the semi-infinite programming we obtain. However, both of them will lead to higher time complexity. In practice, it is enough that we run the FW algorithm until an $\epsilon = 10^{-5}$ optimal solution is obtained in the field of target tracking, since the smaller ϵ does not change the performance of the filter anymore.

V. Properties of the Proposed Filter

In this section, we first show that the proposed nonlinear DSMF is equivalent to the linear SMF proposed by Schweppe [8] if the system becomes linear. Second, the stable analysis of the proposed filter is established.

A. On the Equivalence of the Filters for the Linear Dynamic Systems

When the nonlinear systems in (1) and (2) become linear systems, specifically, let $f_k(\mathbf{x}_k) = \mathbf{F}_k\mathbf{x}_k$ and $h_k(\mathbf{x}_k) = \mathbf{H}_k\mathbf{x}_k$, then we have

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{w}_k \tag{58}$$

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \tag{59}$$

where $\mathbf{F}_k$ and $\mathbf{H}_k$ are the state transition matrix and the measurement matrix, respectively.

In the prediction step of the DSMF, our goal is to derive a minimum size ellipsoid $\mathcal{E}_{f_k}$ containing the nonlinear transformation $\mathcal{F}_k = \{\mathbf{F}_k\mathbf{x}_k, \mathbf{x}_k \in \mathcal{E}_k\}$. However, the linear map of an ellipsoid is an ellipsoid [9]; then, $\mathcal{F}_k = \mathcal{E}_{f_k}$. Thus, the closed-form solution

$$\hat{\mathbf{x}}_{f_k} = \mathbf{F}_k\hat{\mathbf{x}}_k \tag{60}$$

$$\mathbf{P}_{f_k} = \mathbf{F}_k\mathbf{P}_k\mathbf{F}_k^T \tag{61}$$

of the semi-infinite optimization problem (12), (13) can be obtained without sampling, where $\hat{\mathbf{x}}_k$ and $\mathbf{P}_k$ are the center and shape matrix of the ellipsoid $\mathcal{E}_k$, respectively. In fact, Step 2 in Algorithm 1 is ignored. Finally, the prediction step (17), (18) is obtained by replacing $\hat{\mathbf{x}}_{f_k}$ and $\mathbf{P}_{f_k}$ in (60) and (61).

In the measurement update step, when the measurement function is linear, there is no need to use Step 4 in Algorithm 1, which solves the semi-infinite programming problem. Comparing (30) with (59), we can obtain the recursive formula for the measurement update by replacing $\hat{\mathbf{z}}_{k+1}$, $\mathbf{P}_{\mathbf{x}}$, and $\mathbf{P}_{z_{k+1}}$ by $\mathbf{y}_{k+1}$, $\mathbf{H}_{k+1}$, and $\mathbf{R}_{k+1}$ in (31)–(34).

In summary, the proposed DSMF only needs to calculate Steps 3 and 5 in Algorithm 1 without solving the semi-infinite problem by Steps 2 and 4, when the nonlinear system becomes linear. Therefore, this recursive formula of the DSMF is equivalent to the famous SMF algorithm derived by Schweppe [8]. However, in most of the applications, the dynamic systems are nonlinear, which cannot be dealt with by the linear SMF [8].

B. Stability Analysis

In this subsection, inspired by the works of [35] and [36], we prove the asymptotical stability result of the proposed DSMF for some special nonlinear dynamic systems, i.e.,

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{w}_k$$

$$\mathbf{y}_k = h_k(\mathbf{x}_k) + \mathbf{v}_k.$$

Here, the systems have a linear state equation and a nonlinear measurement equation.

First, based on (60) and (61), the prediction step in (17) and (18) can be rewritten as follows:

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}_k\hat{\mathbf{x}}_k \tag{62}$$

$$\hat{\mathbf{P}}_{k+1|k} = \alpha_k\mathbf{F}_k\mathbf{P}_k\mathbf{F}_k^T + \mathbf{Q}_k^*. \tag{63}$$

Second, using Lemma 4 to (33) and setting $\delta_k = 0$, the measurement update step in (31)–(34) is equivalent to

$$\hat{\mathbf{x}}_{k+1} = \hat{\mathbf{x}}_{k+1|k} + \mathbf{K}_{k+1}(\hat{\mathbf{z}}_{k+1} - \mathbf{P}_{\mathbf{x}}\hat{\mathbf{x}}_{k+1|k}) \tag{64}$$

$$\mathbf{P}_{k+1} = (\mathbf{I} - \mathbf{K}_{k+1}\mathbf{P}_{\mathbf{x}})\hat{\mathbf{P}}_{k+1|k} \tag{65}$$

where the gain matrix $\mathbf{K}_{k+1}$, constant α_k, and noise matrices are defined as

$$\mathbf{K}_{k+1} = \hat{\mathbf{P}}_{k+1|k}\mathbf{P}_{\mathbf{x}}^T(\mathbf{P}_{\mathbf{x}}\hat{\mathbf{P}}_{k+1|k}\mathbf{P}_{\mathbf{x}}^T + \mathbf{P}_{\mathbf{z}_{k+1}}^*)^{-1}$$

$$\hat{\mathbf{P}}_{k+1|k} \triangleq \frac{\mathbf{P}_{k+1|k}}{1-\rho_{k+1}}, \alpha_k = \frac{1+p_k^{-1}}{1-\rho_{k+1}} \geq 1$$

$$\mathbf{P}_{\mathbf{z}_{k+1}}^* = \frac{\mathbf{P}_{\mathbf{z}_{k+1}}}{\rho_{k+1}}, \mathbf{Q}_k^* = \frac{(1+p_k)\mathbf{Q}_k}{1-\rho_{k+1}}. \tag{66}$$

Before showing the asymptotical stability of the recursions (62)–(65), some useful assumptions are given.

Assumption 1: The shape matrix of the system noise and the system matrix are bounded by

$$\underline{q} \leq \|\mathbf{Q}_k^*\| \leq \bar{q}$$

$$\underline{z} \leq \|\mathbf{P}_{z_k}^*\| \leq \bar{z}$$

$$\underline{f} \leq \|\mathbf{F}_k\| \leq \bar{f}$$

$$\underline{p} \leq \|\mathbf{P}_{\mathbf{x}}\| \leq \bar{p}.$$

Assumption 2: The discrete time-varying systems (58) and (30) are observable.

Lemma 3 (see [36]): If Assumptions 1 and 2 hold and $\mathbf{P}_0$ is a positive-definite matrix, then there exist some real numbers $\underline{s}, \underline{k}$ and $\bar{s}, \bar{k}$ such that

$$\underline{s}\mathbf{I} \le \mathbf{P}_k \le \bar{s}\mathbf{I}, \underline{s}\mathbf{I} \le \mathbf{P}_{k|k-1} \le \bar{s}\mathbf{I} \tag{67}$$

$$\underline{k} \le \|\mathbf{I} - \mathbf{K}_k\mathbf{P}_{\mathbf{x}}\| \le \bar{k}. \tag{68}$$

Define the estimation error as

$$\zeta_k = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}. \tag{69}$$

Next, the following proposition is used to analyze the stability of the proposed filter.

Proposition 2: Suppose Assumptions 1 and 2 hold; then, the estimation error ζ_k of the DSMF converges to a upper bound ϵ_1 with

$$\epsilon_1 = \frac{\tau_2 + \sqrt{\tau_2^2 + 4\tau_1\tau_3}}{2\tau_1}$$

where $\tau_1 = \frac{1-\alpha_k^{-1}}{\bar{s}} + \frac{1}{\alpha_k \bar{s}^2(\underline{s}^{-1}+\alpha_k \bar{f}^2/\underline{q})}$, $\tau_2 = \frac{2\bar{r}\bar{f}\bar{k}}{\underline{s}}$, $\tau_3 = \frac{2\bar{r}^2}{\underline{s}}$, and $\bar{r}$ is the upper bound of $\mathbf{a}_k = \mathbf{F}_k\mathbf{K}_k\eta_k + \mathbf{w}_k$, i.e., $\|\mathbf{a}_k\| \le \bar{r}$.

Proof: See Appendix C.

VI. Simulations

In this section, we compare the performance among DSMF and UKF [5], ESMF [20], and NSMF [23] by two typical examples, and they show the advantages and effectiveness of the proposed DSMF. In the simulation, the number of samples in the DSMF is $m = 100$.

A. Nonlinear Measurement System

Consider a constant acceleration model of tracking a target in two dimensions [37]. The state contains the position and the velocity of x and y directions. The dynamic system is

$$\mathbf{x}_{k+1} = \mathbf{F}_k\mathbf{x}_k + \mathbf{w}_k \tag{70}$$

$$\mathbf{y}_{k+1} = h_{k+1}(\mathbf{x}_{k+1}) + \mathbf{v}_{k+1} \tag{71}$$

where

$$\mathbf{F}_k = \begin{bmatrix} 1 & 0 & T_0 & 0 \\ 0 & 1 & 0 & T_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$h_k = \begin{bmatrix} \sqrt{(p_k^x - 420)^2 + (p_k^y - 420)^2} \\ \arctan\frac{p_k^y - 420}{p_k^x - 420} \end{bmatrix}.$$

Here, T_0 is the time sampling interval with $T_0 = 1$, $\mathbf{x}_k$ is the state at time k, and $\mathbf{x}_k = [p_k^x, p_k^y, v_k^x, v_k^y]$. The initial state is $\mathbf{x}_0 = [503055]$. Moreover, the process noise $\mathbf{w}_k$ and the measurement noise $\mathbf{v}_k$ are taking value in specified ellipsoidal sets $\mathcal{W}_k$ and $\mathcal{V}_k$, respectively. The shape matrix of the ellipsoidal sets $\mathcal{W}_k$ and $\mathcal{V}_k$ are

$$\mathbf{Q}_k = 10\begin{bmatrix} \frac{T_0^3}{3} & 0 & \frac{T_0^2}{2} & 0 \\ 0 & \frac{T_0^3}{3} & 0 & \frac{T_0^2}{2} \\ \frac{T_0^2}{2} & 0 & T_0 & 0 \\ 0 & \frac{T_0^2}{2} & 0 & T_0 \end{bmatrix}$$

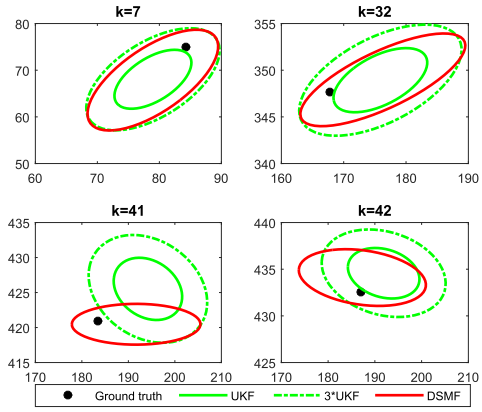

Fig. 4. Estimated ellipsoid by the DSMF and the UKF at different time steps.

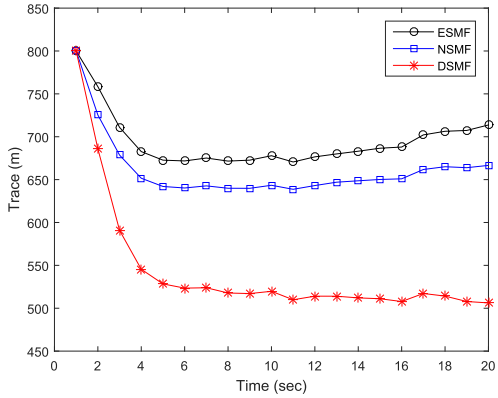

Fig. 5. Size of the estimated ellipsoid by the different methods.

$$\mathbf{R}_k = \begin{bmatrix} 100 & 0 \\ 0 & 0.5 \end{bmatrix}.$$

In the simulation, the noise obeys the uniform distribution. Assume that the initial estimation shape matrix is $\mathbf{P}_0 = 200\mathbf{I}$, and the initial estimation state $\hat{\mathbf{x}}_0$ is a random disturbance around the true state $\mathbf{x}_0$.

The simulation results in this example include three parts: the first part is about the robustness, the second part is about the size of the estimated ellipsoid, and the third part is about the running time.

1) Fig. 4 shows the estimated ellipsoid by the DSMF and the UKF at time steps 7, 32, 41, and 42. The dotted ellipsoid indicates the three times confidence ellipsoid by the UKF. The (stochastic) confidence ellipsoid provided by the UKF is indeed tighter than their deterministic counterparts computed via DSMF, but it does not guarantee the containment of the true state, even if there is three times confidence ellipsoid by the UKF.
2) In Fig. 5, the size $\mathrm{tr}(\mathbf{P}_k)$ of the estimated ellipsoid $\mathcal{E}_k$ is plotted as a function of the time step. It shows that the size of the estimated ellipsoid can quickly converge to a stable value, and the DSMF has the smallest estimated ellipsoid. The reason may be that the DSMF fully extracts the properties of the nonlinear function by solving the

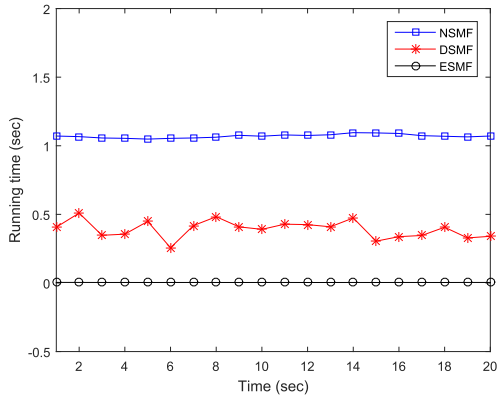

Fig. 6. Running time by the different methods.

semi-infinite optimization problem, rather than linearizing it.

3) The running times of the different SMFs are plotted in Fig. 6, respectively. From Fig. 6, it is shown that the running time of the DSMF stays between the ESMF and the NSMF. The reason is that the ESMF does not need to solve the optimization problem, but the DSMF and the NSMF need to solve some optimization problems by the first-order method and SDP, respectively. In order to get a tradeoff between the performance and the running time, the DSMF method is the best choice.

B. Mobile Robot Localization

Mobile robot location is an important area in artificial intelligence, and one of the key problems about it is the simultaneous localization and mapping (SLAM). In the SLAM problem, mobile robots need maps to locate themselves, and maps need robots to update themselves [38]–[40].

Consider a mobile robot moving through planar environments at time k; the state vector of the robot is defined as $\mathbf{x}_k = [p_k^x, p_k^y, \theta_k]$, where (p_k^x, p_k^y) describes the robot's position in the XY plane, and θ_k is to define the angular orientation. The nonlinear motion model of the robot is given by

$$\mathbf{x}_{k+1} = f_k(\mathbf{x}_k) + \mathbf{w}_k \tag{72}$$

where f_k is a nonlinear state function, which is given as

$$f_k(\mathbf{x}_k) = \begin{bmatrix} p_k^x - \frac{u_p}{u_r}(\sin(\theta_k) - \sin(\theta_k + T_0 u_r)) \\ p_k^y + \frac{u_p}{u_r}(\cos(\theta_k) - \cos(\theta_k + T_0 u_r)) \\ \theta_k + T_0 u_r \end{bmatrix}$$

where $u_p = 0.085$ and $u_r = 0.015$ are the motion command to control the translational velocity and the rotational velocity, respectively. Here, T_0 is the sampling period with $T_0 = 1$. The state noise $\mathbf{w}_k$ in (72) belongs to an ellipsoid $\mathcal{W}_k$, and its shape matrix $\mathbf{Q}_k = \mathrm{diag}(10\mathrm{e}-7, 10\mathrm{e}-7, 10\mathrm{e}-8)$.

Assume that the location of the landmark is $(s^x, s^y) = (50, 50)$; then, the measurement model is [41]

$$\mathbf{y}_k = \begin{bmatrix} r_k \\ \phi_k \end{bmatrix} = \begin{bmatrix} \sqrt{(p_k^x - s^x)^2 + (p_k^y - s^y)^2} \\ \theta_k - \arctan(\frac{p_k^y - s^y}{p_k^x - s^x}) \end{bmatrix} + \mathbf{v}_k. \tag{73}$$

Assume that the measurement noise $\mathbf{v}_k$ is bounded by an ellipsoid $\mathcal{V}_k$, and its shape matrix $\mathbf{R}_k = \mathrm{diag}(1, 1)$.

Indeed, the nonlinear function on the right-hand side of (73) does not exist as a continuous inverse function. However, after some transformations, the following can be obtained:

$$\begin{aligned} \mathbf{P}_{\mathbf{x}}\mathbf{x}_k &= \begin{bmatrix} p_k^x \\ p_k^y \end{bmatrix} = g(\mathbf{y}_k, \mathbf{v}_k, \theta_k) \\ &\triangleq \begin{bmatrix} (\mathbf{y}_k^1 - \mathbf{v}_k^1)\cos(\theta_k - \phi_k - \mathbf{v}_k^2) + s^x \\ (\mathbf{y}_k^1 - \mathbf{v}_k^1)\sin(\theta_k - \phi_k - \mathbf{v}_k^2) + s^y \end{bmatrix} \end{aligned} \tag{74}$$

where $\mathbf{y}_k = [\mathbf{y}_k^1 \mathbf{y}_k^2]$ and $\mathbf{v}_k = [\mathbf{v}_k^1 \mathbf{v}_k^2]$. $\mathbf{P}_{\mathbf{x}}$ is the projection matrix with $\mathbf{P}_{\mathbf{x}} = \left[\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{smallmatrix}\right]$.

Since the nonlinear function g depends not only on $\mathbf{y}_k$ and $\mathbf{v}_k$, but also on θ_k, the measurement ellipsoid cannot be obtained when we only use the information of the measurement $\mathbf{y}_k$ and noise ellipsoid $\mathcal{V}_k$ just as (23). To overcome this, we use the predicted set about θ_k such that $\theta_k \in \mathcal{C}_k^\theta = \{\theta : |\theta - \hat{\theta}_{k|k-1}| \leq \mathbf{P}_{k|k-1}^{3,3}\}$, where $\hat{\theta}_{k|k-1}$ is the third component of the center $\hat{\mathbf{x}}_{k+1|k}$ of the ellipsoid $\mathcal{E}_{k+1|k}$, and $\mathbf{P}_{k|k-1}^{3,3}$ is the third row and the third column of the shape matrix $\mathbf{P}_{k+1|k}$ of the predicted ellipsoid $\mathcal{E}_{k+1|k}$. Then, the measurement ellipsoid $\mathcal{E}(\hat{\mathbf{z}}_k, \mathbf{P}_{z_k})$ can be obtained by solving the following optimization problem:

$$\min\ r(\mathbf{P}_{z_k}) \tag{75}$$

$$\text{s.t. } (\mathbf{z}_k - \hat{\mathbf{z}}_k)^T \mathbf{P}_{z_k}^{-1} (\mathbf{z}_k - \hat{\mathbf{z}}_k) \leq 1 \quad \forall\, \mathbf{z}_k \in \mathcal{C}_k \tag{76}$$

where $\mathcal{C}_k = \{g(\mathbf{y}_k, \mathbf{v}_k, \theta_k) : \mathbf{v}_k \in \mathcal{V}_k, \theta_k \in \mathcal{C}_k^\theta\}$, which is different with the definition in (27).

However, we can also get interior points or boundary points of the set $\mathcal{C}_k$; then, the optimization problem (75), (76) can be solved efficiently by the FW algorithm. Thus, the proposed DSMF can be used to obtain the mobile robot localization.

In our simulation, the initial state $\mathbf{x}_0$ is $[10101]^T$. Moreover, the initial target state estimate $\hat{\mathbf{x}}_0$ is a combination of the actual target state and a random bias vector, and the initial shape matrix $\mathbf{P}_0 = \mathrm{diag}(1, 1, 0.1)$. The root-mean-square error (RMSE) is defined as follows:

$$\mathrm{RMSE}_k = \sqrt{\frac{\sum_{i=1}^{L}(\hat{x}_k^i - x_k^i)^2}{L}} \tag{77}$$

where x_k^i is the true state and $\hat{x}_k^i$ is the state estimation at the ith Monte Carlo; in addition, $L = 50$ in this simulation.

In this example, the RMSE of the state estimation along x and θ directions is plotted as a function of the time steps in Figs. 7 and 8, which show that the RMSE of the DSMF is less than that of the ESMF and the NSMF. The reason may be that the DSMF uses the information of nonlinear functions sufficiently and derives a tighter ellipsoid to cover the state.

VII. Conclusion

In this article, we have proposed the DSMF for a nonlinear dynamic system with unknown but bounded noise. It involves two steps, prediction step and measurement updated step, to recursively compute a bounding ellipsoid to cover the true

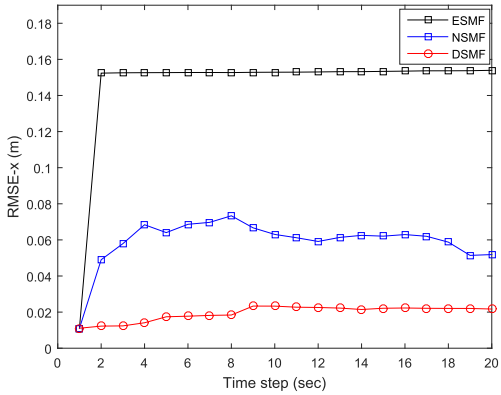

Fig. 7. RMSE of the state p_k^x estimation is plotted as a function of time step.

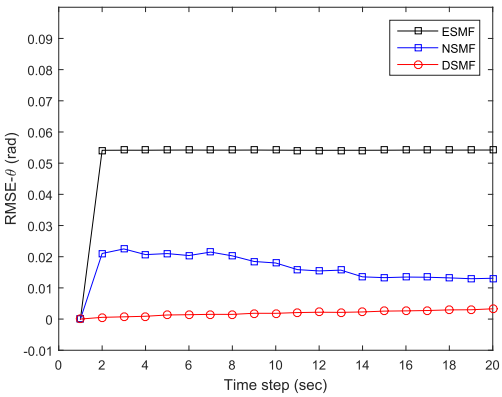

Fig. 8. RMSE of the state θ_k estimation is plotted as a function of time step.

state. The nonlinear system has been translated into the linear system by solving the semi-infinite programming, rather than linearizing the nonlinear function. Moreover, the semi-infinite programming can be solved efficiently by the FW method; then, it is suitable for large-scale problems, and the low computing complexity in each step makes it working online. In addition, the proposed filter is equivalent to the famous SMF derived by Schweppe [8] if the nonlinear dynamic system is degenerated to linear. The stability of the proposed filter is also analyzed for some special nonlinear dynamic systems. Finally, two illustrative examples in the simulations show that the proposed filter performs better. Possible research direction may include multitarget tracking and sensor management in the sensor networks based on the proposed filter. Besides, some interesting open questions are to investigate adaptive filtering for unknown bounds of noises and analyze the stability of the DSMF for the general nonlinear systems.

APPENDIX

A Ellipsoid Containing the Intersection of Two Ellipsoids

It follows that $\mathcal{E}(\hat{\mathbf{x}}_{k+1}, \mathbf{P}_{k+1}) \supseteq \mathcal{E}(\hat{\mathbf{x}}_{k+1|k}, \mathbf{P}_{k+1|k}) \cap \mathcal{E}(\hat{\mathbf{z}}_{k+1}, \mathbf{P}_{z_{k+1}})$ if

$$\begin{aligned}&\mathcal{E}(\hat{\mathbf{x}}_{k+1}, \mathbf{P}_{k+1})\\ &= \{\mathbf{x} : (1-\rho_{k+1})(\mathbf{x}-\hat{\mathbf{x}}_{k+1|k})^T \mathbf{P}_{k+1|k}^{-1}(\mathbf{x}-\hat{\mathbf{x}}_{k+1|k})\\ &+ \rho_{k+1}(\mathbf{P}_{\mathbf{x}}\mathbf{x}-\hat{\mathbf{z}}_{k+1})^T \mathbf{P}_{z_{k+1}}^{-1}(\mathbf{P}_{\mathbf{x}}\mathbf{x}-\hat{\mathbf{z}}_{k+1}) \le 1\} \quad (78)\end{aligned}$$

for any $0 \le \rho_{k+1} \le 1$. Using tedious but straightforward manipulation, we rewrite (78) as

$$\begin{aligned}&\mathcal{E}(\hat{\mathbf{x}}_{k+1}, \mathbf{P}_{k+1})\\ &= \{\mathbf{x} : (\mathbf{x}-\hat{\mathbf{x}}_{k+1})^T \bar{\mathbf{P}}_{k+1}^{-1}(\mathbf{x}-\hat{\mathbf{x}}_{k+1}) \le 1-\delta_{k+1}\} \quad (79)\end{aligned}$$

where $\hat{\mathbf{x}}_{k+1}$, $\bar{\mathbf{P}}_{k+1}$, and δ_{k+1} are defined in (31)–(33), respectively. In addition, $0 \le \delta_{k+1} < 1$. Let $\mathbf{P}_{k+1} = (1-\delta_{k+1})\bar{\mathbf{P}}_{k+1}$; then, we can obtain the final updated ellipsoid as

$$\mathcal{E}(\hat{\mathbf{x}}_{k+1}, \mathbf{P}_{k+1}) = \{\mathbf{x} : (\mathbf{x}-\hat{\mathbf{x}}_{k+1})^T \mathbf{P}_{k+1}^{-1}(\mathbf{x}-\hat{\mathbf{x}}_{k+1}) \le 1\}.$$

Finally, the recursive formula (31)–(34) is obtained.

B Proof for Section IV

The proof of Lemma 1:

Proof: Define an operator $G : \mathbb{R}^{d\times d} \to C(\Omega)$ by

$$[G(\mathbf{H})](\tilde{\mathbf{y}}) \triangleq \tilde{\mathbf{y}}^T \mathbf{H} \tilde{\mathbf{y}} - d \text{on} \Omega$$

where $C(\Omega)$ denotes the Banach space of all continuous on Ω. Thus, the problem (44), (45) can be written as

$$\min \ -\log\det(\mathbf{H}) \quad (80)$$

$$\text{s.t. } G(\mathbf{H}) \le 0. \quad (81)$$

Let us introduce the Lagrangian $L : \mathbb{R}^{d\times d} \times M(\Omega) \to \mathbb{R}$ for problem (80), (81) as

$$L(\mathbf{H}, \mu) = -\log\det(\mathbf{H}) + \langle \mu, G(\mathbf{H}) \rangle \quad (82)$$

where $M(\Omega)$ is the dual space of $C(\Omega)$, and $\langle \cdot, \cdot \rangle$ denotes a duality pair between $C(\Omega)$ and $M(\Omega)$ as

$$\langle \mu, \nu \rangle = \int_\Omega \nu(\tilde{\mathbf{y}}) d\mu.$$

Let us denote the dual operator of the derivative $\nabla G(\mathbf{H})$ [42], which is given by

$$\langle \nabla\mu, G(\mathbf{H}) \rangle = \int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu. \quad (83)$$

By Karush–Kuhn–Tucker conditions, we know that

$$\nabla_H L(\mathbf{H}^*, \mu^*) = -H^{*-1} + \langle \nabla\mu^*, G(\mathbf{H}^*) \rangle = 0 \quad (84)$$

$$\langle \mu^*, \tilde{\mathbf{y}}^T \mathbf{H}^* \tilde{\mathbf{y}} - d \rangle = 0, \mu^* \ge 0. \quad (85)$$

Substituting (83) into (84), the final result can be obtained.

The proof of Lemma 2:

Proof: According to the definition of the Lagrangian function and its gradient in (82) and (84), we conclude that

$$\begin{aligned}\min_{\mathbf{H}} L(\mathbf{H}, \mu) = &\ \log\det\left(\int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu\right)\\ &+ \int_\Omega \left(\tilde{\mathbf{y}}^T \left(\int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu\right)^{-1} \tilde{\mathbf{y}} - d\right) d\mu. \quad (86)\end{aligned}$$

Since

$$\begin{aligned}
&\int_\Omega \tilde{\mathbf{y}}^T \left(\int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu\right)^{-1} \tilde{\mathbf{y}} d\mu \\
&= \int_\Omega \mathrm{tr}\left(\tilde{\mathbf{y}}^T (\int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu)^{-1}\tilde{\mathbf{y}}\right) d\mu \\
&= \int_\Omega \mathrm{tr}\left((\int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu)^{-1}\tilde{\mathbf{y}}\tilde{\mathbf{y}}^T\right) d\mu \\
&= \mathrm{tr}\left((\int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu)^{-1} \int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu\right) \\
&= d
\end{aligned}$$

we can rewrite (86) as

$$\min_{\mathbf{H}} L(\mathbf{H},\mu) = \log\det\left(\int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu\right) + d - d\int_\Omega d\mu. \quad (87)$$

This leads to a first dual problem

$$\max_{\mu} \ \log\det\left(\int_\Omega \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu\right) + d - d\int_\Omega d\mu \quad (88)$$

$$\text{s.t. } \mu \geq 0. \quad (89)$$

Indeed, restricting μ to satisfy $\int_\Omega d\mu = 1$, we obtain the final result.

The proof of Proposition 1:

Proof: Let $\mathbf{H}_y = \mathbf{P}^{-1}$; then, the optimization problem (37), (38) can be rewritten as

$$\min \ -\log\det(\mathbf{H}_y) \quad (90)$$

$$\text{s.t. } (\mathbf{y}-\hat{\mathbf{y}})^T \mathbf{H}_y (\mathbf{y}-\hat{\mathbf{y}}) \leq n \quad \forall\, \mathbf{y} \in \mathcal{C}. \quad (91)$$

Constraint (91) is equivalent to

$$\begin{bmatrix}\mathbf{y}\\1\end{bmatrix}^T \begin{bmatrix}\mathbf{I} & 0\\ -\hat{\mathbf{y}}^T & 1\end{bmatrix} \begin{bmatrix}\mathbf{H}_y & 0\\ 0 & 1\end{bmatrix} \begin{bmatrix}\mathbf{I} & -\hat{\mathbf{y}}\\ 0 & 1\end{bmatrix} \begin{bmatrix}\mathbf{y}\\1\end{bmatrix} \leq d \quad \forall \mathbf{y} \in \mathcal{C}$$

where $d = n+1$. If we define $\tilde{y} = [y^T 1]^T$, it shows that there exists an ellipsoid with shape matrix $\tilde{\mathbf{H}}$

$$\tilde{\mathbf{H}} := \begin{bmatrix}\mathbf{I} & 0\\ -\hat{\mathbf{y}}^T & 1\end{bmatrix} \begin{bmatrix}\mathbf{H}_y & 0\\ 0 & 1\end{bmatrix} \begin{bmatrix}\mathbf{I} & -\hat{\mathbf{y}}\\ 0 & 1\end{bmatrix} \quad (92)$$

which covers the set $\mathcal{C}\times\{1\}$. Note that $-\log\det(\mathbf{H}_y) = -\log\det(\tilde{\mathbf{H}}) \geq -\log\det(\tilde{\mathbf{H}}^*)$, where $\tilde{\mathbf{H}}^*$ is the optimal solution of problem

$$\min -\log\det(\tilde{\mathbf{H}}) \quad (93)$$

$$\text{s.t.} \tilde{\mathbf{y}}^T \tilde{\mathbf{H}} \tilde{\mathbf{y}} \leq d \quad \forall \tilde{\mathbf{y}} \in \mathcal{C}\times\{1\}. \quad (94)$$

From Lemmas 1 and 2, we know that

$$\begin{aligned}
\tilde{\mathbf{H}}^* &= \left(\int \tilde{\mathbf{y}}\tilde{\mathbf{y}}^T d\mu^*\right)^{-1} = \begin{bmatrix}\int \mathbf{y}\mathbf{y}^T d\mu^* & \int \mathbf{y} d\mu^*\\ \int \mathbf{y}^T d\mu^* & 1\end{bmatrix}^{-1} \\
&= \begin{bmatrix}\mathbf{I} & 0\\ -\int \mathbf{y}^T d\mu^* & 1\end{bmatrix} \begin{bmatrix}(\int \mathbf{y}\mathbf{y}^T d\mu^* - \hat{\mathbf{y}}^*(\hat{\mathbf{y}}^*)^T)^{-1} & 0\\ 0 & 1\end{bmatrix} \\
&\quad \cdot \begin{bmatrix}\mathbf{I} & -\int \mathbf{y} d\mu^*\\ 0 & 1\end{bmatrix}. \quad (95)
\end{aligned}$$

Let us set $\mathbf{H}_y^* = (\int \mathbf{y}\mathbf{y}^T d\mu^* - \hat{\mathbf{y}}^*(\hat{\mathbf{y}}^*)^T)^{-1}$ and $\hat{\mathbf{y}}^* = \int \mathbf{y} d\mu^*$; we note that this leads to that $-\log\det(\mathbf{H}_y^*) = -\log\det(\tilde{\mathbf{H}}^*)$, so that

$$-\log\det(\mathbf{H}_y) \geq -\log\det(\mathbf{H}_y^*). \quad (96)$$

Since $\mathbf{H}_y^*$ satisfies constraint (94), we have

$$\tilde{\mathbf{y}}^T \tilde{\mathbf{H}}^* \tilde{\mathbf{y}} \leq d. \quad (97)$$

Substituting (95) into (97) yields

$$(\mathbf{y}-\hat{\mathbf{y}}^*)^T \mathbf{H}_y^* (\mathbf{y}-\hat{\mathbf{y}}^*) \leq n \quad \forall \mathbf{y} \in \mathcal{C} \quad (98)$$

so that the ellipsoid with center $\hat{\mathbf{y}}^*$ and shape matrix $\mathbf{H}_y^*$ contains the set $\mathcal{C}$. According to (96), it proves the minimality of this ellipsoid.

C Proof for Section V

Lemma 4 (see [43]): Let $\mathbf{A}$ and $\mathbf{D}$ be square invertible matrices of sizes $m_A \times m_A$ and $m_D \times m_D$, respectively, and $\mathbf{B}$ and $\mathbf{C}$ be matrices of sizes $m_B \times m_B$ and $m_C \times m_C$, respectively; it follows that

$$(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} = \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1}.$$

The proof of Proposition 2:

Proof: The estimation error is

$$\begin{aligned}
\zeta_{k+1} &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1|k} \\
&= \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k - \mathbf{F}_k \hat{\mathbf{x}}_k \\
&= \mathbf{F}_k(\mathbf{I} - \mathbf{K}_k \mathbf{P}_{\mathbf{x}})\zeta_k + \mathbf{a}_k \quad (99)
\end{aligned}$$

where $\mathbf{a}_k = \mathbf{F}_k \mathbf{K}_k \eta_k + \mathbf{w}_k$; the second equality comes from (58) and (62), and the third equality is due to (30) and (64).

From the notion of $\mathbf{K}_{k+1}$ in (66), we have

$$\mathbf{K}_k(\mathbf{P}_{\mathbf{x}}\hat{\mathbf{P}}_{k|k-1}\mathbf{P}_{\mathbf{x}}^T + \mathbf{P}_{\mathbf{z}_k}^*)\mathbf{K}_k = \hat{\mathbf{P}}_{k|k-1}\mathbf{P}_{\mathbf{x}}^T\mathbf{K}_k.$$

Then, it follows that

$$\begin{aligned}
&(\mathbf{I}-\mathbf{K}_k\mathbf{P}_{\mathbf{x}})\hat{\mathbf{P}}_{k|k-1} \\
&= (\mathbf{I}-\mathbf{K}_k\mathbf{P}_{\mathbf{x}})\hat{\mathbf{P}}_{k|k-1}(\mathbf{I}-\mathbf{K}_k\mathbf{P}_{\mathbf{x}})^T + \mathbf{K}_k\mathbf{P}_{\mathbf{z}_k}^*\mathbf{K}_k^T. \quad (100)
\end{aligned}$$

From (65), we have

$$\mathbf{P}_k \geq (\mathbf{I}-\mathbf{K}_k\mathbf{P}_{\mathbf{x}})\hat{\mathbf{P}}_{k|k-1}(\mathbf{I}-\mathbf{K}_k\mathbf{P}_{\mathbf{x}})^T. \quad (101)$$

Inverting (63), i.e.,

$$\hat{\mathbf{P}}_{k+1|k} = \alpha_k \mathbf{F}_k(\mathbf{P}_k + \alpha_k^{-1}\mathbf{F}_k^{-1}\mathbf{Q}_k^*\mathbf{F}_k^{-T})\mathbf{F}_k^T \quad (102)$$

yields that

$$\hat{\mathbf{P}}_{k+1|k}^{-1} = \alpha_k^{-1}\mathbf{F}_k^{-T}(\mathbf{P}_k + \alpha_k^{-1}\mathbf{F}_k^{-1}\mathbf{Q}_k^*\mathbf{F}_k^{-T})^{-1}\mathbf{F}_k^{-1}. \quad (103)$$

Applying Lemma 4, we obtain

$$\begin{aligned}
&\hat{\mathbf{P}}_{k+1|k}^{-1} \\
&= \alpha_k^{-1}\mathbf{F}_k^{-T}(\mathbf{P}_k^{-1} - \mathbf{P}_k^{-1}(\mathbf{P}_k^{-1} + \alpha_k \mathbf{F}_k \mathbf{Q}_k^{*^{-1}} \mathbf{F}_k^T)^{-1}\mathbf{P}_k^{-1})\mathbf{F}_k^{-1}. \quad (104)
\end{aligned}$$

Using (101) and (65), we obtain

$$\begin{aligned}\hat{\mathbf{P}}_{k+1|k}^{-1} \leq\ & \alpha_k^{-1}\mathbf{F}_k^{-T}(\mathbf{I}-\mathbf{K}_k\mathbf{P}_{\mathbf{x}})^{-T} \\ & \cdot[\hat{\mathbf{P}}_{k|k-1}^{-1}-\hat{\mathbf{P}}_{k|k-1}^{-1}(\mathbf{P}_k^{-1}+\alpha_k\mathbf{F}_k\mathbf{Q}_k^{*^{-1}}\mathbf{F}_k^T)^{-1}\hat{\mathbf{P}}_{k|k-1}^{-1}] \\ & \cdot(\mathbf{I}-\mathbf{K}_k\mathbf{P}_{\mathbf{x}})^{-1}\mathbf{F}_k^{-1}. \end{aligned} \tag{105}$$

To prove the stability of the estimation error, we select the following Lyapunov function:

$$V_k(\zeta_k)=\zeta_k^T\hat{\mathbf{P}}_{k|k-1}^{-1}\zeta_k. \tag{106}$$

Based on Lemma 3, the bound of the Lyapunov function is

$$\frac{\|\zeta_k\|^2}{\bar{s}}\leq V_k(\zeta_k)\leq\frac{\|\zeta_k\|^2}{\underline{s}}. \tag{107}$$

Estimating $V_{k+1}(\zeta_{k+1})$ with (99) and (105), we obtain

$$\begin{aligned}V_{k+1}(\zeta_{k+1})\leq\ & \alpha_k^{-1}\zeta_k^T[\hat{\mathbf{P}}_{k|k-1}^{-1}-\hat{\mathbf{P}}_{k|k-1}^{-1}(\mathbf{P}_k^{-1} \\ & +\alpha_k\mathbf{F}_k\mathbf{Q}_k^{*^{-1}}\mathbf{F}_k^T)^{-1}\hat{\mathbf{P}}_{k|k-1}^{-1}]\zeta_k \\ & +2\mathbf{a}_k^T\hat{\mathbf{P}}_{k+1|k}^{-1}\mathbf{F}_k(\mathbf{I}-\mathbf{K}_k\mathbf{P}_{\mathbf{x}})\zeta_k+\mathbf{a}_k^T\hat{\mathbf{P}}_{k+1|k}^{-1}\mathbf{a}_k. \end{aligned} \tag{108}$$

Applying Lemma 3 and Assumption 1 to (108), we have

$$\begin{aligned}V_{k+1}(\zeta_{k+1})\leq\ & \alpha_k^{-1}\zeta_k^T\hat{\mathbf{P}}_{k|k-1}^{-1}\zeta_k-\frac{\|\zeta_k\|^2}{\alpha_k\bar{s}^2(\underline{s}^{-1}+\alpha_k\bar{f}^2/\underline{q})} \\ & +\frac{2\bar{r}\bar{f}\bar{k}}{\underline{s}}\|\zeta_k\|+\frac{2\bar{r}^2}{\underline{s}}. \end{aligned} \tag{109}$$

According to the notation of τ_1, τ_2, and τ_3, we obtain

$$V_{k+1}(\zeta_{k+1})-V_k(\zeta_k)\leq-\tau_1\|\zeta_k\|^2+\tau_2\|\zeta_k\|+\tau_3. \tag{110}$$

If $\|\zeta_k\|\geq\frac{-\tau_2-\sqrt{\tau_2^2+4\tau_1\tau_3}}{-2\tau_1}$, we have

$$V_{k+1}(\zeta_{k+1})-V_k(\zeta_k)\leq 0. \tag{111}$$

Hence, the estimation error is bounded when the assumptions hold. ■

REFERENCES

[1] J. Lan and X. R. Li, "Multiple conversions of measurements for nonlinear estimation," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4956–4970, Sep. 2017.

[2] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation*. New York, NY, USA: Wiley, 2001.

[3] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Robot. Autom. Mag.*, vol. 13, no. 2, pp. 99–110, Jun. 2006.

[4] E. Masazade, M. Fardad, and P. K. Varshney, "Sparsity-promoting extended Kalman filtering for target tracking in wireless sensor networks," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 845–848, Dec. 2012.

[5] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Nov. 2004.

[6] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[7] B. T. Polyak, S. A. Nazin, C. Durieu, and E. Walter, "Ellipsoidal parameter or state estimation under model uncertainty," *Automatica*, vol. 40, no. 7, pp. 1171–1179, Jul. 2004.

[8] F. C. Schweppe, "Recursive state estimation: Unknown but bounded errors and system inputs," *IEEE Trans. Autom. Control*, vol. AC-13, no. 1, pp. 22–28, Feb. 1968.

[9] A. Kurzhanski and I. Vâlyi, *Ellipsoidal Calculus for Estimation and Control* (Systems and Control: Foundations and Applications). Boston, MA, USA: Birkhäuser, 1997.

[10] H. Wu, W. Wang, and H. Ye, "Model reduction based set-membership filtering with linear state equality constraints," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 2, pp. 1391–1399, Apr. 2013.

[11] S. Rohou, L. Jaulin, L. Mihaylova, and F. L. Bars, "Guaranteed computation of robot trajectories," *Robot. Auton. Syst.*, vol. 97, pp. 76–84, 2017.

[12] L. El Ghaoui and G. Calafiore, "Robust filtering for discrete-time systems with bounded noise and parametric uncertainty," *IEEE Trans. Autom. Control*, vol. 46, no. 7, pp. 1084–1089, Jul. 2001.

[13] X. Shen, Y. Zhu, E. Song, and Y. Luo, "Minimizing euclidian state estimation error for linear uncertain dynamic systems based on multisensor and multi-algorithm fusion," *IEEE Trans. Inf. Theroy*, vol. 57, no. 10, pp. 7131–7146, Oct. 2011.

[14] C. Durieu, E. Walter, and B. Polyak, "Multi-input multi-output ellipsoidal state bounding," *J. Optim. Theory Appl.*, vol. 111, no. 2, pp. 273–303, 2001.

[15] Z. Wang, X. Shen, and Y. Zhu, "Ellipsoidal fusion estimation for multisensor dynamic systems with bounded noises," *IEEE Trans. Autom. Control*, vol. 64, no. 11, pp. 4725–4732, Nov. 2019.

[16] P. H. Leong and G. N. Nair, "Set-membership filtering using random samples," in Proc. *19th Int. Conf. Inf. Fusion*, 2016, pp. 1087–1094.

[17] F. Gu, Y. He, and J. Han, "Active persistent localization of a three-dimensional moving target under set-membership uncertainty description through cooperation of multiple mobile robots," *IEEE Trans. Ind. Electron.*, vol. 62, no. 8, pp. 4958–4971, Aug. 2015.

[18] L. Sun, H. Alkhatib, B. Kargoll, V. Kreinovich, and I. Neumann, "A new Kalman filter model for nonlinear systems based on ellipsoidal bounding," 2018, *arXiv:1802.02970*.

[19] B. Chen and G. Hu, "Nonlinear state estimation under bounded noises," *Automatica*, vol. 98, pp. 159–168, 2018.

[20] E. Scholte and M. E. Campbell, "A nonlinear set-membership filter for on-line applications," *Int. J. Robust Nonlinear Control*, vol. 13, no. 15, pp. 1337–1358, 2003.

[21] R. E. Moore, *Interval Analysis*. Englewood Cliffs, NJ, USA:Prentice-Hall, 1966.

[22] F. Yang and Y. Li, "Set-membership fuzzy filtering for nonlinear discrete-time systems," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 1, pp. 116–124, Feb. 2010.

[23] G. Calafiore, "Reliable localization using set-valued nonlinear filters," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 35, no. 2, pp. 189–197, Mar. 2005.

[24] Y. Liu and X. R. Li, "Measure of nonlinearity for estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2377–2388, May 2015.

[25] L. Ma, Z. Wang, H.-K. Lam, and N. Kyriakoulis, "Distributed event-based set-membership filtering for a class of nonlinear systems with sensor saturations over sensor networks," *IEEE Trans. Cybern.*, vol. 47, no. 11, pp. 3772–3783, Nov. 2017.

[26] S. D. Ahipasaoglu, P. Sun, and M. J. Todd, "Linear convergence of a modified Frank-Wolfe algorithm for computing minimum-volume enclosing ellipsoids," *Optim. Methods Softw.*, vol. 23, no. 1, pp. 5–19, 2008.

[27] M. J. Todd, *Minimum-Volume Ellipsoid: Theory and Algorithms*, vol. 23. Philadelphia, PA, USA: SIAM, 2016.

[28] Z. Wang, X. Shen, Y. Zhu, and J. Pan, "A tighter set-membership filter for some nonlinear dynamic systems," *IEEE Access*, vol. 6, pp. 25351–25362, 2018.

[29] W. Pu, Y.-F. Liu, J. Yan, H. Liu, and Z.-Q. Luo, "Optimal estimation of sensor biases for asynchronous multi-sensor data fusion," *Math. Program.*, vol. 170, no. 1, pp. 357–386, 2018.

[30] G. Calafiore and M. Campi, "Uncertain convex programs: Randomized solutions and confidence levels," *Math. Program.*, vol. 102, pp. 25–46, 2005.

[31] M. A. Armstrong, Basic Topology. New York, NY, USA: Springer-Verlag, 1983.

[32] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory (Studies in Applied Mathematics)*. Philadelphia, PA, USA: SIAM, Jun. 1994.

[33] R. Reemtsen and J. J. Rückmann, *Semi-Infinite Programming*, vol. 25. New York, NY, USA: Springer, 1998.

[34] G. Still, "Discretization in semi-infinite programming: The rate of convergence," *Math. Program.*, vol. 91, no. 1, pp. 53–69, 2001.

[35] Y. Wang, J. Huang, D. Wu, Z.-H. Guan, and Y.-W. Wang, "Set-membership filtering with incomplete observations," *Inf. Sci.*, vol. 517, pp. 37–51, 2020.

[36] K. Reif and R. Unbehauen, "The extended Kalman filter as an exponential observer for nonlinear systems," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2324–2328, Aug. 1999.
[37] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1333–1364, Oct. 2003.
[38] L. Jaulin, "A nonlinear set membership approach for the localization and map building of underwater robots," *IEEE Trans. Robot.*, vol. 25, no. 1, pp. 88–98, Feb. 2009.
[39] B. Chen, G. Hu, D. W. Ho, and L. Yu, "A new approach to linear/nonlinear distributed fusion estimation problem," *IEEE Trans. Autom. Control*, vol. 64, no. 3, pp. 1301–1308, Mar. 2019.
[40] W. Yu, E. Zamora, and A. Soria, "Ellipsoid SLAM: A novel set membership method for simultaneous localization and mapping," *Auton. Robots*, vol. 40, pp. 125–137, 2016.
[41] S. Thrun, "Probabilistic robotics," *Commun. ACM*, vol. 45, no. 3, pp. 52–57, 2002.
[42] S. Ito, Y. Liu, and K. L. Teo, "A dual parametrization method for convex semi-infinite programming," *Ann. Oper. Res.*, vol. 98, nos. 1–4, pp. 189–213, 2000.
[43] R. A. Horn and C. R. Johnson, *Matrix Analysis* 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2012.

Zhiguo Wang received the Ph.D. degree from Sichuan University, Chengdu, China, in 2018.

From 2018 to 2020, he was Postdoctoral Researcher with The Chinese University of Hong Kong, Shenzhen, China. Since 2020, he has been working as a specially appointed Associate Research Fellow with the College of Mathematics, Sichuan University. His current research interests include information fusion, target tracking, machine learning, nonconvex optimization, and analysis and control of uncertain systems.

Xiaojing Shen received the Graduate and Ph.D. degrees from Sichuan University, Chengdu, China, in 2003 and 2009, respectively, and the M.S. degree from Jilin University, Changchun, China, in 2006.

From 2009 to 2011, he was a Postdoctoral Researcher and Assistant Professor with the School of Computer Science, Sichuan University. Since 2013, he has been a Professor with the School of Mathematics, Sichuan University. From 2012 to 2013, he was a Postdoctoral Research Associate with Syracuse University, Syracuse, NY, USA. He is one of the authors of two books entitled *Networked Multisensor Decision and Estimation Fusion: Based on Advanced Mathematical Methods* (Boca Raton, FL, USA: CRC Press, 2012) and *Nonlinear Estimation and Applications to Industrial Systems Control* (G. Rigatos, Ed., ch. 3, pp. 61–88, Hauppauge, NY, USA: Nova Science, 2012). He has also served as a cooperator to several major research institutes. His current research interests include multisensor decision and estimation fusion, multitarget data association and tracking, optimization theory with applications to information fusion, and statistical theory with applications to information fusion.

Dr. Shen is the recipient of the China National Excellent Doctoral Dissertation Award (the highest honor of Ph.D. degree).

Haiqi Liu received the B.S. degree in 2016 from the School of Mathematics, Sichuan University, Chengdu, China, where he is currently working toward the Ph.D. degree.

His current research interests include multitarget tracking and data association, target tracking, and optimization theory with applications to information fusion.

Fanqin Meng received the Graduate degree from the Shandong University of Science and Technology, Qingdao, China, in 2008, and the M.S. and Ph.D. degrees from Sichuan University, Chengdu, China, in 2015 and 2019, respectively.

Since 2019, he has been a Lecturer with School of Automation and Information Engineering, Sichuan University of Science and Technology, Chengdu, China, and a Postdoctoral Researcher with the School of Aeronautics and Astronautics, Sichuan University. His research interests include information fusion, target tracking, localization, machine learning, and convex optimization.

Yunmin Zhu received the B.S. degree from the Department of Mathematics and Mechanics, Beijing University, Beijing, China, in 1968.

From 1968 to 1978, he was with Luoyang Tractor Factory, Luoyang, Henan, China, as a Steel Worker and a Machine Engineer. From 1981 to 1994, he was with the Institute of Mathematical Sciences, Chengdu Institute of Computer Applications, Chengdu Branch, Academia Sinica. Since 1995, he has been a Professor with the Department of Mathematics, Sichuan University, Chengdu, China. From 1986 to 1987, 1989 to 1990, 1993 to 1996, 1998 to 1999, and 2001 to 2002, he was a Visiting Associate or Visiting Professor with the Lefschetz Centre for Dynamical Systems and the Division of Applied Mathematics, Brown University; Department of Electrical Engineering, McGill University; Communications Research Laboratory, McMaster University; and the Department of Electrical Engineering, University of New Orleans. He is the author or co-author of more than 80 papers in international and Chinese journals. He is the author of the books entitled *Multisensor Decision and Estimation Fusion* (Norwell, MA, USA: Kluwer, 2002), *Networked Multisensor Decision and Estimation Fusion: Based on Advanced Mathematical Methods* (Boca Raton, FL, USA: CRC Press, 2012), and *Multisensor Distributed Statistical Decision* (Beijing, China: Science Press, 2000) and a co-author of *Stochastic Approximations* (Shanghai, China: Shanghai Scientific and Technical Publishers, 1996). His research interests include stochastic approximations, adaptive filtering, other stochastic recursive algorithms and their applications in estimations, optimizations, and decisions for dynamic system as well as for signal processing, information compression. In particular, his present major interest is multisensor distributed estimation and decision fusion.

Prof. Zhu is on the Editorial Boards of the *Journal of Systems Science and Complexity* and the *International Journal of Systems Science and Applied Mathematics*.